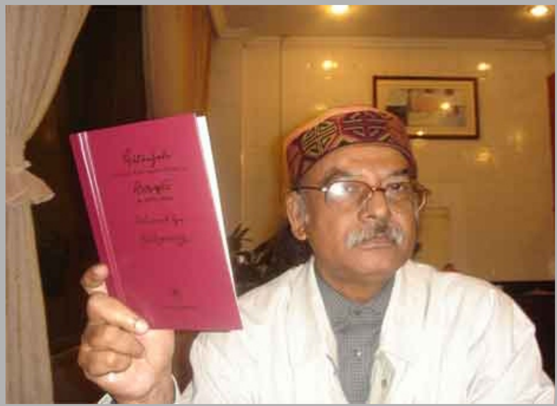


গণিত শিল্পী ছাড়া ‘আকাশ কুসুম’- এর
চাষ-আবাদ করার ক্ষমতা ও দুঃসাহস বিজ্ঞান জগতের
আর কোন সাধকের থাকতে পারে ?

ব্যবহারযোগ্যতা গণিতের কোন মূল ধর্ম নয় । ..... এই অর্থে গণিত
অনেক কাছাকাছি শিল্পের অথবা কবিতার । সুন্দরের যদি কোন সত্য
থাকে তাহলে সে সত্য গণিতেরও ।

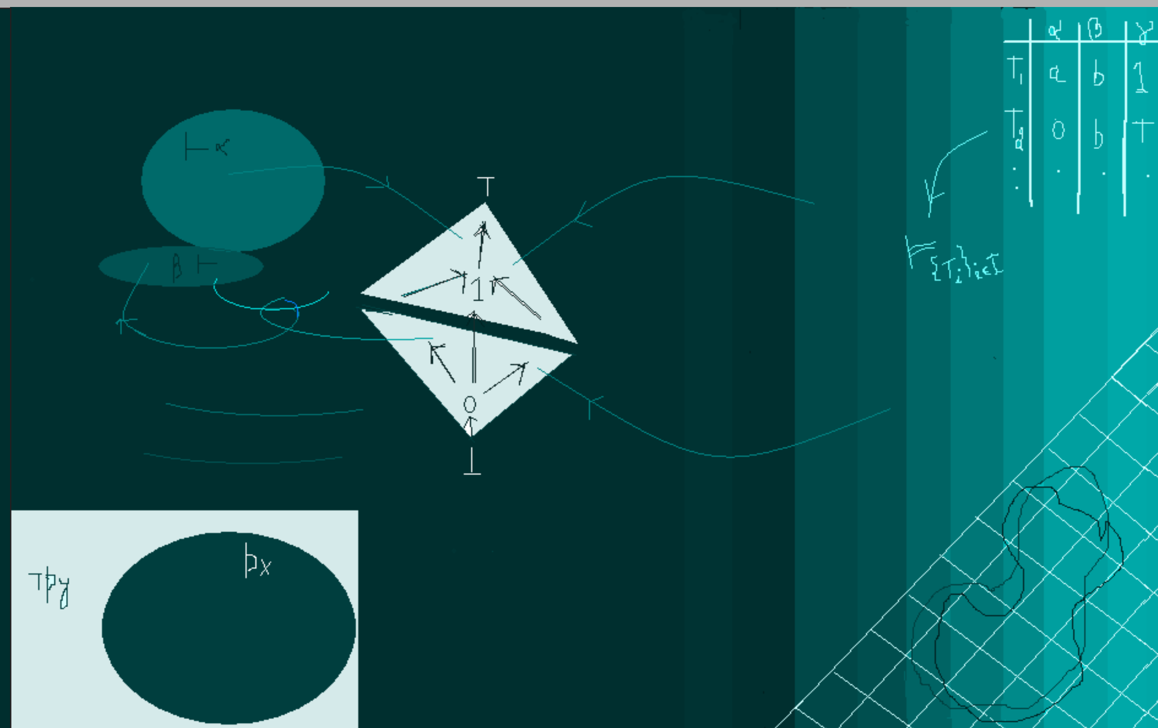
....তার উপরে গড়ে উঠেছে বিভিন্ন গণিত ।
নানা ধরনের লজিকও প্রস্তাবিত হচ্ছে । ...... আপাতত
শত ফুল বিকশিত হোক সাথে কিছু আগাছাও

Calcutta Logic Circle 2011

Essays in honour of Mihir Kr. Chakraborty

CLC
2011

Reminiscing Ideas and Interactions



Essays in honour of Mihir Kr. Chakraborty

Reminiscing

Ideas and Interactions

Reminiscing

Ideas and Interactions

Essays in honour of
Mihir Kr. Chakraborty

Calcutta Logic Circle
2011

• **CLC 2011**

Printed by Dutta Printer, 5 Raja S.C. Mullick Road, Kolkata-700032, India

Cover designed by Soma Dutta and Supriya Pahari

Quotations on the back cover are from the following articles in books of Bengali essays written by Mihir Kr. Chakraborty.

"গণিত শিল্পী ছাড়া ...": Mihir Chakraborty, Ākāśkusumēr Adhikār, in *Gaṇitēr Dhārāpāt Ō Galpasalpa*, Nandimukh Samsad, Kolkata, India, 2005, p. 15.

"ব্যবহারযোগ্যতা গণিতের কোন ...": Mihir Chakraborty, Sundarēr Satya, Gaṇitēr Satya, in *Amimāṃsār Ālō-Āñdhāri*, Nandimukh Samsad, Kolkata, India, 2010, p. 111-112.

"...তার উপরে গড়ে ...": Mihir Chakraborty, Gaṇitjagatē Parbaparibartan, in *Gaṇitēr Dhārāpāt Ō Galpasalpa*, Nandimukh Samsad, Kolkata, 2005, p. 33.

Sketch on the right by Mihir Kr. Chakraborty (courtesy: Soma Dutta)


16/9/05

KEEP THE FLAG FLYING

(To Professor Mihir K. Chakraborty: An Appreciation)

Haragauri N. Gupta

I was very pleased to receive an invitation to write for a volume of essays in honour of Professor Mihir K. Chakraborty. I have had the privilege of knowing Professor Chakraborty for over thirty five years and take this opportunity to send an appreciation of his work in Logic and his service in the area of Logical research.

Mihir K. Chakraborty's contributions in Logic are considerable. His dedication to Logic is an abiding source of inspiration for his students and has resulted in the creation of the Calcutta Logic Circle (CLC) some twenty five years ago. Members of the circle belong to diverse disciplines. Under Professor Chakraborty's leadership they gather together to discuss problems arising in their fields, specifically problems of a logical nature. "Logic has been esteemed as the lamp of all studies, the source of all actions and the shelter of all virtues" to quote the ancient Indian savant Kautilya of the third century BC. Modern Logic provides a space where Arts (Humanities), Sciences and Technology interplay to the mutual benefit of them all. It also gives valuable insights for Philosophy, Mathematics, Computer Science, Engineering, and Linguistics, to name only a few. The growth of Logic in recent times is astounding. For about two thousand years since Aristotle (384-324 BC) Logic remained in the shape which Aristotle gave to it. Then, about the middle of the 19th century some mathematicians (Boole, De Morgan, Peirce, and Schroeder et al) began to investigate the logical processes as they existed at the time. They brought to the open the mathematical nature of much that went in the name of Logic. For example, the propositional core of logic was seen to constitute a mathematical system, ushering in the Boolean Algebra. The Aristotelian System of syllogisms was extended to become the Predicate Calculus of today (Frege, Russell, Hilbert). (Why did this much needed generalization have to wait for so long?) Fundamental questions on the newly emerging logical systems were raised and answered using mathematical methods. This was possible because of a parallel event in Mathematics itself, namely, the emergence of Set Theory (Cantor, Dedekind, Zermelo). The concept of truth (at least in formalized languages) was cleared up by Alfred Tarski (1901-1983) putting Logic of derived knowledge on reliable semantic foundations. Model Theory emerged. Notions of completeness, decidability and axiomatizability of theories (mathematical) began to engage the attention of mathematicians. Kurt Godel (1906-1978) made groundbreaking discoveries in Logic and Set Theory. His arithmetization of logical syntax can be perceived as fulfilment of the Leibnizian dream. Way back in the seventeenth century G.W. Leibniz (1646-1716) had advanced the idea that Logic ought to be seen as a variety of Mathematics both in content and in form (calculemos). All this happened in a brief span of time (1840-1930) blurring the distinction between Mathematics and Logic. Logic is no longer in the exclusive care of philosophers. Doors of cross-fertilization with other disciplines have been flung open. Alternative, non-standard logics have emerged (many-valued logics, modal logics, Logic of fuzzy sets, Logic of rough sets). Computer Science appeared and is growing in rapid strides. We are witness to a magnificent era of cooperation, wondering at the power that interplay of disciplines can generate.

Members of the Calcutta Logic Circle are held together by this shared sense of wonder. They welcome everything that can be seen as the logical heart of their concerns. They meet regularly to discuss logical matters. Under Mihir K. Chakraborty's leadership, the Calcutta Logic Circle has brought about a renaissance of logical studies in Calcutta and beyond. The present volume of essays is a fitting testimonial to the measure of respect and gratefulness which his students, colleagues and members of the Calcutta Logic Circle feel toward him. It is a lovely way indeed to felicitate a loving and caring teacher.

From overseas, I send Mihir Chakraborty warmest good wishes on his sixty fifth birthday and wish him many more years of service to Logic and Learning. To the Calcutta Logic Circle I say:

Keep the Flag Flying

Haragauri N. Gupta

Emeritus Professor,
Department of Mathematics & Statistics,
University of Regina,
Canada.

guptahn2000@yahoo.ca

Contents

Preface

It was in 1985 when a mathematician in love with logic invited one of his teachers Haragauri Narayan Gupta, a prominent student of Alfred Tarski, to deliver a course of lectures designed for charting a roadmap for serious studies in logic at his workplace, the University of Calcutta, Kolkata. A motley collection of people also interested in logic, but from various disciplines, gathered to go through the course. The course ended, but the craving for more and more of what was served was felt by the mathematician, and, encouraged by the mood of the others around him, he took up the challenging task of forming a community of logic lovers single handedly, providing the infrastructure, the planning and also the resources in terms of materials and seminars, all by himself. The community, although informal, bonded slowly into a close knit group, the cementing factor being the gentle yet irresistible personality of Mihir Kumar Chakraborty, the mathematician we have been talking about.

The community, called Calcutta Logic Circle, is still alive, informally. From initially being a study group, it has now become a research group gaining the affection, attention and support of many around the country and abroad. In its journey many have joined, and some have drifted away, but a core group – not all of them residing in Kolkata – remains, keeping in touch with all those who care for it. The central figure, around whom all in the group assembled sometime or other, called variously as MKC, Mihirda, Mihirbabu, Mihir, Sir, etc. (depending on the context and relationship), turns sixty five this year. It is not only in a show of respect for this person – an extra-ordinary teacher – but also in a show of solidarity with the ideas, vision and efforts of his in letting this community live, that Calcutta Logic Circle presents a collection of essays on the occasion (prepared as a surprise for him).

A variety of articles, touching various aspects of his life and philosophy, and ranging from technical articles on different areas of logic and mathematics that he has worked on, to very personal ones, written by his students, friends, colleagues and members of Calcutta Logic Circle, give this volume a unique look. We hope that, through these articles, the readers would get a glimpse of his vision and ideas, as well as of him as a person. A few of his graduate students have presented outlines of their ongoing doctoral work, thus giving an idea of his current research projects as well.

We are indebted to Haragauri Narayan Gupta for supporting our effort wholeheartedly and contributing to this volume. Our sincere thanks go to all the other authors, who have unhesitatingly given their consent to contribute, and then valuable time and effort in preparing the articles. The members of Calcutta Logic Circle also thank their family, friends and colleagues for providing the much needed support in making this volume a reality.

Calcutta Logic Circle, 2011

http://home.iitk.ac.in/~mohua/CLC.html

Algebraic Studies in Pawlak's Rough Set Theory

Mohua Banerjee

1 Introduction

Indiscernibility has been at the focus of debates, for ages. I ventured to make it the focus of my Ph.D. thesis [1] as well, spurred on by my supervisor's very infectious enthusiasm[1] about the notion. We explored it from different points of view, which led to the consideration of different *approximate identities.* In particular, we considered the situation when there was lack of complete information on the objects of the domain involved – which is just what Pawlak's *rough set theory* [31] deals with. The theory hinges on the notion of an *approximation space* (U, R), where R is an indiscernibility relation on the domain of discourse U, represented by a partition.

Since the interest in the thesis was on foundational issues, we naturally took a close look at operations on rough sets. Pawlak had, of course, meticulously listed some immediate properties, in his first paper itself, and later in his book [32]. However, that was clearly only a starting point, and not surprisingly, soon after, questions on algebras that could be formed with rough sets as basic entities surfaced, and started being addressed by many. We joined the bandwagon, so to speak, but soon realized that we must put the study in perspective. The result turned out to be a list of algebras, both new and well-known, the common feature being that each has a 'rough set' as an instance. In this article, we present an account of the developments in the study of algebraic structures that result only from 'classical' rough set theory, i.e. ones using notions defined by Pawlak. There have also been algebraic investigations in various generalizations of rough set theory, but covering those is beyond the scope of this article. A brief presentation of these studies is made at the end in [6]. The account presented here is a re-organized and updated version of that in [6]. We also refer to [23].

In order to develop an algebra an instance (or model) of which would be rough sets, it is clearly necessary to specify the definition of a rough set. Interestingly, we find that there are four set-based definitions of a (Pawlak) rough set itself, viz. Definitions 3, 4, 5, 6 below. Apart from these, we also have the 'operator-oriented' view [37] reflected by Definitions 1, 2.

First, a brief on the preliminaries [32]. Consider an approximation space (U, R). Let $[x]$ denote the equivalence class of the element x of U under the equivalence relation R. For any subset A of U, the *lower approximation* $\underline{A}$ of A in the approximation space (U, R) is the set $\{x \in U : [x] \subseteq U\}$, i.e. the union of equivalence classes contained in A. The *upper approximation* $\overline{A}$ of A in

[1] This is only one amongst various other subjects, mostly non-standard, many "non-academic", that Prof. Mihir Chakraborty continues to inspire me about, in quite a similar manner.

(U,R) is the set $\{x \in U : [x] \cap A \neq \emptyset\}$: the union of equivalence classes properly intersecting A. The *boundary* BnA of A is the set $\overline{A} \setminus \underline{A}$ – it consists of elements possibly, but not definitely, in A. A *definable set* in (U,R) is one whose boundary is empty.
Two sets A, B in (U,R) are said to be *roughly equal*, denoted $A \approx B$, provided $\underline{A} = \underline{B}$, as well as $\overline{A} = \overline{B}$. $\approx$, clearly is an equivalence relation on $\mathcal{P}(U)$, the power set of U.

Definition 1. *[32] $A \subseteq U$ is a rough set in (U,R), provided $Bn(A) \neq \emptyset$.*

For generality's sake however, we could remove the restriction in the definition above, and term *any* subset A of U rough. A definable set then becomes a special case of a rough set. Moreover, to keep the context clear, we could have

Definition 2. *The triple (U,R,A) is called a rough set [3].*

Definition 3. *(cf. [23]) The pair $(\underline{A}, \overline{A})$, for each $A \subseteq U$, is called a rough set in (U,R).*

Definition 4. *[29] The pair $(\underline{A}, \overline{A}^c)$, for each $A \subseteq U$, is called a rough set in (U,R), where ^{c} denotes complementation in $\mathcal{P}(U)$.*

Definition 5. *[27] Given an approximation space (U,R), a rough set is an ordered quadruple (U,R,L,B), where (i) L, B are disjoint subsets of U, (ii) both L and B are definable sets in (U,R), and (iii) for each $x \in B$ there exists $y \in B$ such that $x \neq y$ and xRy (i.e. no equivalence class contained in B is a singleton).*

Definition 6. *[31] A rough set in (U,R) is an equivalence class of $\mathcal{P}(U)/\approx$.*

Let us fix an approximation space (U,R). We denote by $\mathcal{F}_i,\ i = 1,\ldots,5$, the collection of all rough sets corresponding to (U,R), given by Definition 1, 3, 4, 5, 6 respectively, only the version of Definition 5 where (U,R) is fixed in all the quadruples, is considered to get $\mathcal{F}_4$. We observe that Definitions 3, 4, 5 and 6 are equivalent to each other for any given (U,R), in the sense that there is a one-one correspondence between families $\mathcal{F}_2,\ \mathcal{F}_3,\ \mathcal{F}_4$ and $\mathcal{F}_5$. Indeed, for any $A \subseteq U$, the entities $(\underline{A}, \overline{A})$, $(\underline{A}, \overline{A}^c)$ and the equivalence class $[A]$ of A in $\mathcal{P}(U)/\approx$ are identifiable. Again, for a fixed (U,R), a quadruple (U,R,L,B) is essentially the pair (L,B), and due to condition (iii) of Definition 5, one can always find a subset A of U such that $\underline{A} = L$ and $Bn(A) = B$. Hence Definition 5 may be reformulated as follows: the pair $(\underline{A}, Bn(A))$ for each $A \subseteq U$ is a rough set so long as (U,R) remains unchanged. So, via this interpretation, Definition 5 also becomes equivalent to 3, 4 and 6. These definitions have been said to represent the 'set-oriented' view of rough sets, viz. when rough sets are defined through pairs of definable sets. In contrast, Definitions 1 and 2 are said to represent the 'operator-oriented' view, where the term 'rough' is used as an adjective for the subset A of U, and *operators* of lower and upper approximations are defined on U, in order to approximately describe A.

However, as we shall see in this article, these equivalent definitions yield different (though related) algebras, by taking different definitions of union, intersection, complementation and other algebraic operations.

Apart from the above families $\mathcal{F}_1, \ldots, \mathcal{F}_5$, another collection that makes an appearance in the algebraic studies is $\mathcal{R} := \{(D_1, D_2),\ D_1 \subseteq D_2,\ D_1, D_2 \in \mathcal{D}\}$, where $\mathcal{D}$ denotes the collection of all definable sets of a fixed approximation space (U, R).
Notice that $\mathcal{R}$ is a generalization of $\mathcal{F}_2 :\ \mathcal{F}_2 \subseteq \mathcal{R}$. Further, since $D_2 \setminus D_1$ may include singleton equivalence classes, $\mathcal{F}_2$ may be a proper subset of $\mathcal{R}$.

A summary of the structures obtained by considering the families above are presented in Section 2. Section 3 presents the algebras obtained from the operator-based definitions, that include ones coming from *information systems*, another basic notion in rough set theory. Both Sections 2 and 3 give the main results proved about the algebras as well. In Section 4, we outline the various relationships that are obtained among these algebras. Section 5 concludes the article.

2 Algebras from set-based definitions

One finds that the different algebras emerging from the set-based definitions of rough sets are instances of :

1. quasi-Boolean algebras [19, 33];
2. (a) topological quasi-Boolean algebras [4, 36];
   (b) pre-rough algebras [5];
   (c) rough algebras [5];
3. regular double Stone algebras [15, 29];
4. complete atomic Stone algebras [12];
5. semi-simple Nelson algebras [29];
6. 3-valued Łukasiewicz algebras [29, 18].

In the sequel, we sketch how exactly these algebras come about, starting from the definitions of rough sets. Broadly, the scheme is of the following nature. As remarked in the Introduction, the primary task is to fix the definition of a rough set and therefore the corresponding family. The next step is to take an appropriate operation of union / intersection / complementation / interior etc. on the family, to give rise to a class of algebraic structures, say $\mathcal{RS}$. This leads, on abstraction (according to the properties of the operations in $\mathcal{RS}$), to one of the classes of algebras in the preceding list. In many cases, the connection between $\mathcal{RS}$ and the corresponding class (say $\mathcal{A}$) of abstract algebras, is formalized by establishing a *representation* result. One demonstrates a correspondence $c : \mathcal{A} \to \mathcal{RS}$ such that any element $A \in \mathcal{A}$ is *isomorphic* to a subalgebra of $c(A)$. In some cases, a reverse representation is proved.

We note that when the family $\mathcal{F}_2$, or the more general $\mathcal{R}$ is considered, a natural definition for the operations of union and intersection of the members would be the following.

Definition 7.

– $(D_1, D_2) \sqcup (D_1', D_2') \;:=\; (D_1 \cup D_1', D_2 \cup D_2')$,
– $(D_1, D_2) \sqcap (D_1', D_2') \;:=\; (D_1 \cap D_1', D_2 \cap D_2')$.

A restriction of these operations to the subclass $\mathcal{F}_2$ would give:

Definition 8.

– $(\underline{A}, \overline{A}) \sqcup (\underline{B}, \overline{B}) \;:=\; (\underline{A} \cup \underline{B}, \overline{A} \cup \overline{B})$,
– $(\underline{A}, \overline{A}) \sqcap (\underline{B}, \overline{B}) \;:=\; (\underline{A} \cap \underline{B}, \overline{A} \cap \overline{B})$.

One needs to ensure closure here, i.e. to check whether the right-hand entities in the above do belong to $\mathcal{F}_2$. This issue has been addressed by authors in [12, 6, 24]. For the sake of completeness, let us present a summary of the results here. The question is, given an approximation space (U, R), whether there are subsets C, D of U such that
$\underline{A} \cup \underline{B} \;=\; \underline{C}$, $\overline{A} \cup \overline{B}$ $(= \overline{A \cup B}) \;=\; \overline{C}$, and
$\underline{A} \cap \underline{B}$ $(= \underline{A \cap B}) \;=\; \underline{D}$, $\overline{A} \cap \overline{B} = \overline{D}$.
We give three pairs (C_i, D_i), $i = 1, 2, 3$, of such sets, which then clearly must be coordinate-wise roughly equal.

$\boldsymbol{(C_1, D_1)}$ [6] Include in C_1, one of the sets, say A, the lower approximation of the other ($\underline{B}$), and add all the elements of $A \cup B$ that are in $Bn(A \cup B)$. In other words,
$C_1 := A \cup \underline{B} \cup ((A \cup B) \cap Bn(A \cup B))$.
To obtain D_1, we add to $A \cap B$, all elements of A (or B) that lie in the equivalence classes $[x]$ which are within $\overline{A} \cap \overline{B}$, but outside $\overline{A \cap B}$. So
$D_1 := (A \cap B) \cup (A \cap ((\overline{A} \cap \overline{B}) \setminus (\overline{A \cap B})))$.

$\boldsymbol{(C_2, D_2)}$ Another way to come up with the sets C and D is as follows [12].

Definition 9. *An upper sample P of A in (U, R) is a subset of U such that $P \subseteq A$ and $\overline{P} = \overline{A}$. An upper sample P is minimal, if there is no upper sample Z of A with $Z \subsetneq P$.*

Let P be a minimal upper sample of $\overline{A} \cup \overline{B}$. We take
$C_2 := \underline{A} \cup \underline{B} \cup P$, and
$D_2 := (\underline{A} \cap \underline{B}) \cup P$.

$\boldsymbol{(C_3, D_3)}$ Yet another approach [24] is to consider a set B_0 roughly equal to B:

$$B_0 := \underline{B} \cup (B \cap \overline{A}^c) \cup (\overline{B} \cap A \setminus \underline{A}) \cup (B \cap \underline{A}).$$

The idea behind the construction of B_0 is to take any equivalence class in $\overline{B}$ and account for all its possible relationships with the set A. Define
$C_3 := A \cup B_0$, and
$D_3 := A \cap B_0$.
One can then show that $\underline{A} \cup \underline{B} = \underline{A \cup B_0}$ and $\overline{A} \cap \overline{B} = \overline{A \cap B_0}$.

2.1 Quasi-Boolean algebras

Definition 10. *[35] A* quasi-Boolean algebra *(or a* De Morgan lattice*) is a bounded distributive lattice* $(A, \leq, \vee, \wedge, 0, 1)$ *with a unary operation* $\neg$ *that satisfies involution (*$\neg\neg a = a$*, for each* $a \in A$*), and makes the De Morgan identities hold.*

Iwiński [19] and Pomykała [33] show that rough sets form structures that are quasi-Boolean algebras. Iwiński, presenting a 'rough algebra' for the first time, follows Definition 3. The general collection $\mathcal{R}$ instead of $\mathcal{F}_2$ is considered. Operations of join ($\sqcup$) and meet ($\sqcap$) on $\mathcal{R}$ are given by Definition 7. It may be noted that any definable set A of (U, R) is identifiable with the pair (A, A) of $\mathcal{R}$. Further,

Definition 11. $\neg(D_1, D_2) := ({D_2}^c, {D_1}^c)$*, where* ^{c} *denotes ordinary set-theoretic complementation.*

$\neg$ satisfies the De Morgan identities, and when restricted to definable sets, is the usual complement. But it does not satisfy the laws of Boolean complementation in general.

Proposition 1. *[19]* $(\mathcal{R}, \sqcup, \sqcap, \neg, 0, 1)$ *is a complete atomic quasi-Boolean algebra, where* $0 := (\emptyset, \emptyset)$ *and* $1 := (U, U)$*. Atoms are of the form* $(\emptyset, A)$*,* A *being an elementary set of* (U, R)*. The definable sets form a maximal Boolean subalgebra of* $(\mathcal{R}, \sqcup, \sqcap, \neg, 0, 1)$*.*

Does the converse of this proposition hold? Let us see.
A basic finite quasi-Boolean algebra is $\mathcal{U}_0 := (\{0, a, b, 1\}, \vee, \wedge, \neg, 1)$. It is a diamond as a lattice, viz.

$$\begin{tikzcd}[column sep=small] & \mathbf{1} & \\ \mathbf{a} \arrow[ur] & & \mathbf{b} \arrow[ul] \\ & \mathbf{0} \arrow[ul] \arrow[ur] & \end{tikzcd}$$

and $\neg$ is given by the equations :

$$\neg 0 := 1,\ \neg 1 := 0,\ \neg a := a,\ \neg b := b.$$

It is known [35] that any quasi-Boolean algebra is isomorphic to a subalgebra of the product $\prod_{i \in I} \mathcal{U}_i$, where I is a set of indices, and $\mathcal{U}_i = \mathcal{U}_0$.

Hence, to address the converse of proposition 1, it seems natural to ask if $\mathcal{U}_0$ is isomorphic to some $\mathcal{R}$. The answer is in the negative, since for any member (D_1, D_2) of $\mathcal{R}$ other than $(\emptyset, X)$, $\neg(D_1, D_2) \neq (D_1, D_2)$, whereas in $\mathcal{U}_0$, $\neg a = a,\ \neg b = b$ and $a \neq b$. So the class $\mathcal{R}$ is a proper subclass of the class of quasi-Boolean algebras.

No representation result is proved in [19]. However, if one considers the family $\mathcal{F}_2$, such a result is obtained in [33].

Clearly, $(\mathcal{F}_2, \sqcup, \sqcap, \neg, 0, 1)$ is also a quasi-Boolean algebra, the operations being restrictions of those in $\mathcal{R}$. But [33] says more. An important notion involved here is that of an 'individual atom' – a singleton elementary class. Let us denote by S, the collection of all individual atoms in the approximation space $< X, R >$. Two simple examples of quasi-Boolean algebras are the two and three element chains $\mathcal{B}_0 := (\{0,1\}, \vee, \wedge, \neg, 1)$ and $\mathcal{C}_0 := (\{0,a,1\}, \vee, \wedge, \neg, 1)$ respectively. $\neg$ is defined as: $\neg 0 := 1,\ \neg 1 := 0,\ \neg a := a$.

Theorem 1. *$(\mathcal{F}_2, \sqcup, \sqcap, \neg, 0, 1)$ is isomorphic to a subalgebra of the product $\prod_{i \in I} \mathcal{U}_i$, where I is a set of indices, and $\mathcal{U}_i = \mathcal{B}_0$ or $\mathcal{U}_i = \mathcal{C}_0$, for each $i \in I$.*

It should be mentioned that Pomykała came up with a number of algebraic structures that have $\mathcal{F}_2$ as domain. These differ from each other with respect to the complementation and implication operations chosen.

2.2 Topological quasi-Boolean algebras

Definition 12. *[4, 36] A* topological quasi-Boolean algebra *(tqBa) is a quasi-Boolean algebra $(A, \leq, \vee, \wedge, \neg, 0, 1)$ with an* interior *(unary) operation L that satisfies*

L1 $La \leq a$,
L2 $L(a \wedge b) = La \wedge Lb$,
L3 $LLa = La$,
L4 $L1 = 1$ *and*
L5 $MLa = La$,

where M is the closure *operation, viz.* $Ma := \neg L \neg a, a, b \in A$.

Proceeding from Section 2.1, one may define an interior operation L on $(\mathcal{R}, \sqcup, \sqcap, \neg, 0, 1)$:

Definition 13. $L(D_1, D_2) := (D_1, D_1),\ D_1, D_2 \in \mathcal{D}$.
Thus, the closure M on $\mathcal{R}$ is given by $M(D_1, D_2) = (D_2, D_2)$.

On the other hand, one may start from Definition 6, and define, for $[A], [B] \in \mathcal{P}(U)/\approx$,

– $[A] \sqcap [B] := [A \sqcap B]$,
– $\neg[A] := [A^c]$,
– $L[A] := [\underline{A}]$,

where $A \sqcap B$ is the set D_1 given in the beginning of this section.

One then obtains

Proposition 2. *L as in Definition 13 gives the tqBa $(\mathcal{R}, \sqcup, \sqcap, \neg, L, 0, 1)$. Restricting L to $\mathcal{F}_2$ makes $(\mathcal{F}_2, \sqcup, \sqcap, \neg, 0, 1)$ form a tqBa. The quotient set $\mathcal{P}(U)/\approx$ yields a tqBa structure as well, with the preceding definitions of $\sqcap, \neg$ and L.*

The tqBa on $\mathcal{P}(U)/\approx$ is isomorphic to that on $\mathcal{F}_2$. It is also isomorphic to that on $\mathcal{R}$, provided of course, no definable set in (U, R) is a singleton.

No representation result of rough structures with respect to tqBa's has been proved. As a matter of fact, the class of tqBa's itself is open to investigation. Algebraically, the following is the only known result so far.

Proposition 3. *[8] TqBa's form a variety that is not a discriminator variety.*

The tqBas on $\mathcal{P}(U)/\approx$ and $\mathcal{F}_2$ satisfy more properties, as we shall see in Sections 2.3 and 2.4.

2.3 Pre-rough algebras

The following are added to the definition of a tqBa to get a pre-rough algebra.

Definition 14. *[5] A* pre-rough algebra *is a tqBa* $(A, \leq, \vee, \wedge, \neg, L, 0, 1)$ *in which*

- $\neg La \vee La = 1$,
- $L(a \vee b) = La \vee Lb$,
- $La \leq Lb, Ma \leq Mb$ *imply* $a \leq b$.

One may define an 'implication' operation in this structure as

$$a \Rightarrow b := (\neg La \sqcup Lb) \sqcap (\neg Ma \sqcup Mb).$$

Observation 1 *In a pre-rough algebra* $\mathcal{P} := (A, \leq, \vee, \wedge, \neg, L, 0, 1)$*,* $\mathcal{L}(A) := (L(A), \leq, \vee, \wedge, \neg, 0, 1)$*, where* $L(A) := \{La : a \in A\}$*, is a Boolean algebra (using the first axiom in Definition 14). It may also be noticed that* $L(A) = M(A)$*.*

Example 1. Let $\mathcal{T} := (A, \leq, \vee, \wedge, \neg, L, 0, 1)$, where $A := \{0, a, 1\}$,

$$\begin{array}{c} 1 \\ \uparrow \\ a \\ \uparrow \\ 0 \end{array}$$

with $\neg 0 := 1$, $\neg a := a$, $\neg 1 := 0$, $L0 := 0, La := 0$, $L1 := 1$. $\mathcal{T}$ is the smallest non-trivial pre-rough algebra.

The tqBa's on $\mathcal{P}(U)/\approx$ and $\mathcal{F}_2$ (and also on $\mathcal{R}$) are pre-rough algebras.

A representation result [5] shows that any pre-rough algebra is, in fact, an algebra of pairs of Boolean elements.

Theorem 2. *Any pre-rough algebra* $(A, \leq, \sqcap, \sqcup, \neg, L, 0, 1)$ *is isomorphic to the pre-rough algebra formed by the set* $B := \{(La, Ma) : a \in A\}$*. The operations on* B *are defined just by abstracting those on* $\mathcal{F}_2$*.*

2.4 Rough algebras

Definition 15. *[5] A* rough algebra $\mathcal{P} := (A, \leq, \sqcap, \sqcup, \neg, L, \Rightarrow, 0, 1)$ *is a pre-rough algebra such that the subalgebra* $(L(A), \leq, \sqcap, \sqcup, \neg, 0, 1)$ *of* $\mathcal{P}$ *is*

- *complete and*
- *completely distributive, i.e.* $\sqcup_{i \in I} \sqcap_{j \in J} a_{i,j} = \sqcap_{f \in J^I} \sqcup_{i \in I} a_{i,f(i)}$*, for any index sets* I, J *and elements* $a_{i,j}, i \in I, j \in J$*, of* $L(A)$*,* J^I *being the set of maps of* I *into* J*.*

The pre-rough algebras on each of $\mathcal{P}(U)/\approx$, $\mathcal{F}_2$ and $\mathcal{R}$, are rough algebras as well. The following representation result is then obtained.

Theorem 3. *Any rough algebra is isomorphic to a subalgebra of* $(\mathcal{R}, \sqcup, \sqcap, \neg, L, 0, 1)$ *corresponding to some approximation space* (U, R).

In fact, one can show [6] that

Corollary 1. *Any rough algebra is isomorphic to a subalgebra of* $\mathcal{P}(U')/\approx$ *for some approximation space* (U', R').

2.5 Complete atomic Stone algebras

Definition 16. *A* Stone algebra *is a bounded distributive lattice* $(A, \leq, \vee, \wedge, 0, 1)$ *which has a* pseudo-complement * *on* A, *i.e.* $y \leq x^*$ *if and only if* $y \wedge x = 0$, *and which satisfies the* Stone identity, *viz.* $x^* \vee x^{**} = 1$.

In [34], Pomykała defines * on $\mathcal{F}_2$ as:

Definition 17. $(\underline{A}, \overline{A})^* := (\overline{A}^c, \overline{A}^c)$, $(\underline{A}, \overline{A}) \in \mathcal{F}_2$.

Then one obtains (with $\sqcup, \sqcap$ as in Definition 8, and $0, 1$ as in Proposition 1)

Proposition 4. $(\mathcal{F}_2, \sqcup, \sqcap, ^*, 0, 1)$ *is a Stone algebra.*

However, no representation is obtained.

Starting from Definition 6, [12] arrives at an enhanced rough structure. $(\mathcal{P}(U)/\approx, \leq)$ is clearly a partially ordered set, $\leq$ being defined in terms of rough inclusion, i.e. $[A] \leq [B]$, if and only if A is roughly included in B, $[A], [B] \in \mathcal{P}(U)/\approx$. Now operations of join ($\cup_\approx$), meet ($\cap_\approx$) on $\mathcal{P}(U)/\approx$ and ('exterior') complementation ($^{\mathrm{ex}}$) are defined.

For a subset A of U, an *upper sample* P is such that $P \subseteq A$ and $\overline{P} = \overline{A}$. An upper sample P of A is *minimal*, if there is no upper sample Z of A with $Z \subsetneq P$. Then

Definition 18.
- $[A] \cup_\approx [B] := [\underline{A} \cup \underline{B} \cup P]$, *where* P *is a minimal upper sample of* $\overline{A} \cup \overline{B}$, *and*
- $[A] \cap_\approx [B] := [(\underline{A} \cap \underline{B}) \cup P]$, *where* P *is a minimal upper sample of* $\overline{A} \cap \overline{B}$.
- $[A]^{\mathrm{ex}} := [(\overline{A})^c]$.

One may note that $\emptyset$ is included among elementary sets. For a finite domain U,

Proposition 5. $(\mathcal{P}(U)/\approx, \cup_\approx, \cap_\approx, ^{\mathrm{ex}}, [\emptyset], [U])$ *is a complete atomic Stone algebra, where the atoms are determined by proper subsets of the elementary sets or by singleton elementary sets in* (U, R).

Again, no representation is obtained. Such a result is found though, on introducing a further operation on the family of rough sets $\mathcal{F}_2$.

2.6 Regular double Stone algebras

Definition 19. *A double Stone algebra (dSa) is a Stone algebra* $(A, \vee, \wedge, ^*, 0, 1)$ *which has a* dual pseudo-complement $^+$, *i.e.*

– $y \geq x^+$ *if and only if* $y \vee x = 1$,

and which satisfies

– $x^+ \wedge x^{++} = 0$.

The dSa is regular *if, for all* $x, y \in A$,

– $x \wedge x^+ \leq y \vee y^*$ *holds.*

(This is equivalent to requiring that

– $x^* = y^*$, $x^+ = y^+$ *imply* $x = y$, *for all* $x, y \in A$.*)*

[15] introduces a dual pseudo-complementation $^+$ on $\mathcal{F}_2$ and gets, further to Proposition 4,

Proposition 6. $(\mathcal{F}_2, \sqcup, \sqcap, ^*, ^+, 0, 1)$, *for a given approximation space* (U, R), *is a regular dSa, where* $(\underline{A}, \overline{A})^+ := (\underline{A}^c, \underline{A}^c)$.

As a representation result, Comer obtains

Theorem 4. *Any regular dSa is isomorphic to a subalgebra of* $(\mathcal{F}_2, \sqcup, \sqcap, ^*, ^+, 0, 1)$ *for some approximation space* (U, R).

2.7 Semi-simple Nelson algebras

Definition 20. *[35] A* Nelson algebra *is a quasi-Boolean algebra* $(A, \wedge, \vee, \neg, 0, 1)$ *equipped with a unary operation* $\sim$ *and a binary operation* $\rightarrow$ *such that, for any* $a, b, x \in A$,

– $a \wedge \neg a \leq b \vee \neg b$,
– $a \wedge x \leq \neg a \vee b$ *if and only if* $x \leq a \rightarrow b$,
– $a \rightarrow (b \rightarrow c) = (a \wedge b) \rightarrow c$,
– $\sim a = a \rightarrow \neg a = a \rightarrow 0$.

A Nelson algebra A *is* semi-simple, *if* $a \vee \sim a = 1$, *for all* $a \in A$.

$\neg$ and $\sim$ are the 'strong' and 'weak' negations on A respectively.

These algebras are discussed in the context of rough sets in [29], which considers finite domains, and adopts Definition 4. It is observed that

Proposition 7. $(\mathcal{F}_3, \sqcap, \sqcup, \neg, \sim, \rightarrow, 0, 1)$ *is a semi-simple Nelson algebra, the operations being defined as:*

– $(\underline{A_1}, \overline{A_1}^c) \sqcap (\underline{A_2}, \overline{A_2}^c) := (\underline{A_1} \cap \underline{A_2}, \overline{A_1}^c \cup \overline{A_2}^c)$,

- $(\underline{A_1}, \overline{A_1}^c) \sqcup (\underline{A_2}, \overline{A_2}^c) \ := \ (\underline{A_1} \cup \underline{A_2}, \overline{A_1}^c \cap \overline{A_2}^c)$,
- $(\underline{A_1}, \overline{A_1}^c) \rightarrow (\underline{A_2}, \overline{A_2}^c) \ := \ (\underline{A_1}^c \cup \underline{A_2}, \underline{A_1} \cap \overline{A_2}^c)$,
- $\neg(\underline{A_1}, \overline{A_1}^c) \ := \ (\overline{A_1}^c, \underline{A_1})$, *and*
- $\sim (\underline{A_1}, \overline{A_1}^c) \ := \ (\underline{A_1}^c, \underline{A_1})$.

The representation theorem is as follows.

Theorem 5. *Any finite semi-simple Nelson algebra is isomorphic to* $(\mathcal{F}_3, \sqcap, \sqcup, \neg, \sim, \rightarrow, 0, 1)$ *for some approximation space* (U, R).

$\mathcal{F}_3$ also forms a Stone as well as a regular double Stone algebra with suitable operations. The operations are derived from those which make $\mathcal{F}_3$ a Nelson algebra (cf. proposition 7). $\sqcap, \sqcup$ remain the same, while the pseudo-complementation * is taken as $\neg \sim \neg$, and the dual pseudo-complementation $^+$ as $\sim$.

2.8 3-valued Łukasiewicz algebras

Definition 21. *(cf. [11]) A 3-valued Łukasiewicz (Moisil) algebra* $(A, \leq, \wedge, \vee, \neg, M, 0, 1)$ *is such that* $(A, \leq, \wedge, \vee, \neg, 0, 1)$ *is a quasi-Boolean algebra and* M *is a unary operator on* A *satisfying, for all* $a, b \in A$

- $M(a \wedge b) = Ma \wedge Mb$,
- $M(a \vee b) = Ma \vee Mb$,
- $Ma \wedge \neg Ma = 0$,
- $MMa = Ma$,
- $M\neg Ma = \neg Ma$,
- $\neg M \neg a \leq Ma$, *and*
- $Ma = Mb, M\neg a = M\neg b$ *imply* $a = b$.

A direct representation result concerning this class of algebras has been obtained in [18]. However, the same has been concluded through relationships of 3-valued Łukasiewicz algebras with other algebras, in both [29] and [2]. We shall elaborate on this in Section 4.

[18] considers Definition 3, i.e. the family $\mathcal{F}_2$, and also finite domains.

With the operator M as in Definition 13 (restricted to $\mathcal{F}_2$), $\sqcup, \sqcap$ as in Definition 8, $\neg$ as in Definition 11, $0 := (\emptyset, \emptyset)$, $1 := (U, U)$ (cf. Proposition 1), one finds that

Proposition 8. $(\mathcal{F}_2, \sqcup, \sqcap, \neg, M, 0, 1)$ *is a 3-valued Łukasiewicz algebra.*

The representation theorem is as follows.

Theorem 6. *Every 3-valued Łukasiewicz algebra is isomorphic to a subalgebra of* $(\mathcal{F}_2, \sqcup, \sqcap, \neg, M, 0, 1)$ *corresponding to some approximation space* (U, R).

2.9 Other algebras

In the special situation when the approximation space has no singleton elementary sets in it, [29] observes that $\mathcal{F}_2$ with its pairs in reverse order, viz. the collection of pairs $(\overline{A}, \underline{A})$, $A \subseteq U$, turns out to be a *Post algebra* of order three [35]. Therefore [13], it is a 3-valued Łukasiewicz algebra with a centre (i.e. an element c such that $\neg c = c$).

In the general situation (with no restriction on the approximation space), the same structure can be made into an algebra that is a generalization of a Post algebra, viz. a certain chain-based lattice of order three [16].

3 Algebras from operator-based definitions

As mentioned in the Introduction, in the operator-oriented approach, the lower and upper approximations are viewed as unary operators on the domain. In the simplest manifestation here, we see that rough set structures form instances of Boolean algebras with operators.

3.1 Boolean algebras with operators

Definition 22. *(cf. [10]) A* Boolean algebra with operators *is a Boolean algebra* $(A, \vee, \wedge, \sim, 0, 1,)$ *along with a collection* $\{f_i\}_{i \in I}$ *of operators on A, where I is an index set. Each n-ary operator* $f_i, i \in I$, *satisfies the two following properties:*

O1. $f_i(a_1, ..., 0, ..., a_n) = 0$ *(*normality*) and*

O2. $f_i(a_1, ..., a_i \vee a_{i'}, ..., a_n) = f_i(a_1, ..., a_i, ..., a_n) \vee f_i(a_1, ..., a_{i'}, ..., a_n)$ *(*additivity*).*

An instance of a Boolean algebra with operators, is a *topological Boolean algebra*:

Definition 23. *[35] A* topological Boolean algebra *(tBa) is a Boolean algebra* $(A, \vee, \wedge, \sim, 0, 1)$*, that has a unary operation L satisfying the properties* L1-L4 *of an interior given in Definition 12.*

If (U, R) is an approximation space, the set $\mathcal{P}(U)$ of all subsets of U forms a *topological Boolean algebra*. The interior operation on $\mathcal{P}(U)$ is nothing but the lower approximation with respect to the approximation space (U, R), regarded as a unary operator on U, i.e. $L(X) := \underline{X}$, for any $X \subseteq U$. L, in fact, satisfies all the properties L1-L5 of Definition 12. The upper approximation operator, denoted by $^-$ say, would then satisfy the properties of a closure operator, which include $O1, O2$ of Definition 22. In other words, as observed in [18], $(\mathcal{P}(U), \cup, \cap, ^c, ^-, \emptyset, U)$ forms a *monadic Boolean algebra* [25], that is an instance of a Boolean algebra with the single binary operator $^-$. In [38], the lower approximation operator $_-$, dual to $^-$, is also brought into the picture. Yao studies an abstraction of $(\mathcal{P}(U), \cup, \cap, ^c, _-, ^-)$, and considers a structure $(\mathcal{P}(U), \cup, \cap, ^c, L, H)$, called a *rough set algebra*. $(\mathcal{P}(U), \cup, \cap, ^c, H)$ is first taken to be simply a Boolean algebra with the operator H, and L as an operation on $\mathcal{P}(U)$ dual to H. The result of imposing further conditions on the two operators are then investigated.

3.2 Algebras from complete information systems

A basic notion in rough set theory is that of an *information system* [31]: a table providing information about the values that objects of a domain are assigned under a given set of *attributes*.

Definition 24. *A* complete information system *(CIS)* $\mathcal{S} := (U, \mathcal{A}, \bigcup_{a\in\mathcal{A}} \mathcal{V}_a, f)$*, comprises a non-empty set* U *of* objects*,* $\mathcal{A}$ *of* attributes*,* $\mathcal{V}_a$ *of* attribute-values *for each* $a \in \mathcal{A}$*, and* information function $f : U \times \mathcal{A} \rightarrow \bigcup_{a\in\mathcal{A}} \mathcal{V}_a$ *such that* $f(x, a) \in \mathcal{V}_a$*.*

Any subset B of the attribute set $\mathcal{A}$ induces an *indiscernibility relation* $Ind_B^{\mathcal{S}}$ (an equivalence relation on U):

$$(x, y) \in Ind_B^{\mathcal{S}}, \text{ if and only if } f(x, a) = f(y, a) \text{ for all } a \in B.$$

Each relation $Ind_B^{\mathcal{S}}$ induces an *upper approximation operator* $\overline{Ind_B^{\mathcal{S}}}$ on $\mathcal{P}(U)$, mapping a set $X(\subseteq U)$ to $\overline{X}_{Ind_B^{\mathcal{S}}}$. Similarly, we have a *lower approximation operator* $\underline{Ind_B^{\mathcal{S}}}$, for each $B \subseteq \mathcal{A}$.

3.2.1 Cylindric algebras

The structure $\mathcal{B}_{\mathcal{S}} := (\mathcal{P}(U), \cup, \cap, ^c, \{\overline{Ind_B^{\mathcal{S}}}\}_{B\subseteq\mathcal{A}}, \emptyset, U)$, is called a *knowledge approximation algebra of type* $\mathcal{A}$ *derived from the CIS* $\mathcal{S}$ in [14]. For each $B \subseteq \mathcal{A}$, the structure $(\mathcal{P}(U), \cup, \cap, ^c, \overline{Ind_B^{\mathcal{S}}}, \emptyset, U)$ is called the (upper) *approximation closure algebra of* B. It may be noted that this is an instance of a monadic Boolean algebra.

Now $(\mathcal{P}(U), \cup, \cap, ^c, \emptyset, U)$ is not only a Boolean algebra, as mentioned in Section 3.1, it is also complete, and atomic. In fact, Comer observes that a knowledge approximation algebra of type $\mathcal{A}$ is an instance of a general algebraic structure, an *abstract knowledge approximation algebra of type* $\mathcal{A}$.

Definition 25. *An* abstract knowledge approximation algebra of type $\mathcal{A}$ *consists of a complete atomic Boolean algebra* $(S, \vee, \wedge, \sim, 0, 1)$*, and a family of functions* $K_B : S \rightarrow S$*,* $B \subseteq \mathcal{A}$*,* $\mathcal{A}$ *being a finite set. Moreover, the functions satisfy the following, for* $x, y \in S$ *and* $B, C \subseteq \mathcal{A}$*.*

- $K_B(0) = 0$*.*
- $K_B(x) \geq x$*.*
- $K_B(x \wedge K_B(y)) = K_B(x) \wedge K_B(y)$*.*
- *If* $x \neq 0$ *then* $K_\emptyset(x) = 1$*.*
- $K_{B\cup C}(x) = K_B(x) \wedge K_C(x)$*, if* x *is an atom of* S*.*

Comer then indicates the relation of approximation closure algebras with *cylindric algebras*.

Definition 26. *A* cylindric algebra of dimension $|\mathcal{A}|$ *[17] is a structure* $\mathfrak{A} := (U, \wedge, \neg, 0, \{\Lambda_a\}_{a\in\mathcal{A}}, \{\mu_{(a,b)}\}_{(a,b)\in\mathcal{A}\times\mathcal{A}})$*, where* $(U, \wedge, \neg, 0)$ *is a Boolean algebra, and* Λ_a*,* $\mu_{(a,b)}$ *are respectively unary and nullary operations on* U*, such that*

(L_1) *$\Lambda_a(0) = 0$,*
(L_2) *$x \leq \Lambda_a(x)$,*
(L_3) *$\Lambda_a(x \wedge \Lambda_a(y)) = \Lambda_a(x) \wedge \Lambda_a(y)$,*
(L_4) *$\Lambda_a(\Lambda_b(x)) = \Lambda_b(\Lambda_a(x))$,*
(L_5) *$\mu_{(a,a)} = 1$,*
(L_6) *If $a \neq b, c$, then $\mu_{(b,c)} = \Lambda_a(\mu_{(b,a)} \wedge \mu_{(a,c)})$,*
(L_7) *If $a \neq b$, then $\Lambda_a(\mu_{(a,b)} \wedge x) \wedge \Lambda_a(\mu_{(a,b)} \wedge \neg x) = 0$.*

We then see that

Proposition 9. *Approximation closure algebras are complete atomic cylindric algebras of dimension one.*

A representation theorem is subsequently obtained.

Theorem 7. *Every complete atomic cylindric algebra of dimension one is isomorphic to an approximation closure algebra. In fact, every cylindric algebra of dimension one is embeddable in an approximation closure algebra.*

3.2.2 CIS-algebras

The knowledge approximation algebra derived from a CIS does not give a complete description of an information system: attribute and attribute-value pairs, which are the salient features of an information system, do not appear in this description. Thus these algebras are unable to capture the fact that approximations are *induced* in information systems by attributes and their values. In a recent work [22], a class of algebras that takes care of this aspect, has been studied in detail. We give the basic definitions and summary of the results obtained.

In a CIS $\mathcal{S} := (U, \mathcal{A}, \bigcup_{a\in\mathcal{A}} \mathcal{V}_a, f)$, we observe that each *descriptor* [32] (a, v), where a is an attribute, and v an attribute-value, also determines a nullary operation (constant) $c^{\mathcal{S}}_{(a,v)}$ on $\mathcal{P}(U)$:

$$c^{\mathcal{S}}_{(a,v)} := \{x \in U : f(x, a) = v\}.$$

Let Ω be the tuple $(\mathcal{A}, \mathcal{V})$, for a fixed $\mathcal{A}$ and $\mathcal{V} := \bigcup_{a\in\mathcal{A}} \mathcal{V}_a$, and let $\mathcal{D}$ denote the set of all descriptors obtained from Ω.

Definition 27. *A* complete information system algebra (in brief, *CIS-algebra*) of type Ω generated by a complete information system $\mathcal{S} := (U, \mathcal{A}, \mathcal{V}, f)$ *is the structure*

$$\mathcal{S}^* := (\mathcal{P}(U), \cap, ^c, \{\underline{Ind^{\mathcal{S}}_B}\}_{B\subseteq\mathcal{A}}, \{c^{\mathcal{S}}_\gamma\}_{\gamma\in\mathcal{D}}, \emptyset).$$

So a CIS-algebra generated by a CIS $\mathcal{S}$ is an extension of the knowledge approximation algebra derived from $\mathcal{S}$ with a collection of nullary operations.

Making use of some properties that actually turn out to be characterizing properties of CIS-algebras, a notion of an *abstract CIS-algebra* is defined.

Definition 28. *An* abstract CIS-algebra of type Ω *is a tuple*

$$\mathfrak{A} := (U, \wedge, \neg, \{L_B\}_{B\subseteq\mathcal{A}}, \{d_\gamma\}_{\gamma\in\mathcal{D}}, 0),$$

where $(U, \wedge, \neg, 0)$ *is a Boolean algebra and* L_B *and* d_γ *are respectively unary and nullary (constant) operations on* U *satisfying the following:*

(C_1) $\bigvee_{v\in\mathcal{V}_a} d_{(a,v)} = 1$;
(C_2) $d_{(a,v)} \wedge d_{(a,u)} = 0$ *when* $v \neq u$;
(C_3) $L_C(x) \leq L_B(x)$ *for* $C \subseteq B \subseteq \mathcal{A}$;
(C_4) $d_{(a,v)} \leq L_{\{a\}}(d_{(a,v)})$;
(C_5) $d_{(b,v)} \wedge L_{B\cup\{b\}}(x) \leq L_B(\neg d_{(b,v)} \vee x)$;
(C_6) $L_\emptyset(x) \neq 0$ *implies* $x = 1$.

The CIS-algebra $\mathcal{S}^*$ generated by a CIS $\mathcal{S}$ is an abstract CIS-algebra. On the other hand, an abstract CIS-algebra $\mathfrak{A}$ determines a unique CIS $\mathfrak{A}_* := (PF(\mathfrak{A}), \mathcal{A}, \mathcal{V}, f_\mathfrak{A})$, where $f_\mathfrak{A} : PF(\mathfrak{A}) \times \mathcal{A} \to \mathcal{V}$ is the map defined as

$$f_\mathfrak{A}(\Gamma, a) := v \text{ if and only if } d_{(a,v)} \in \Gamma.$$

We also note that the reduct $(U, \wedge, \neg, \{L_B\}_{B\subseteq\mathcal{A}}, 0)$ of $\mathfrak{A}$ determines a *complex algebra* [10] as follows. Let $PF(\mathfrak{A})$ denote the set of all prime filters of $\mathfrak{A}$.
For each $B \subseteq \mathcal{A}$, let us consider the binary relation $Q_B^\mathfrak{A} \subseteq PF(\mathfrak{A}) \times PF(\mathfrak{A})$:

$$(\Gamma, \Delta) \in Q_B^\mathfrak{A} \text{ if and only if } L_B(x) \in \Gamma \text{ implies } x \in \Delta.$$

The relations $Q_B^\mathfrak{A}$ are used to define the operators $m_B^\mathfrak{A} : \mathcal{P}(PF(\mathfrak{A})) \to \mathcal{P}(PF(\mathfrak{A}))$:

$$m_B^\mathfrak{A}(X) := \{\Gamma \in PF(\mathfrak{A}) : \text{ for all } \Delta \text{ such that } (\Gamma, \Delta) \in Q_B^\mathfrak{A},\ \Delta \in X\}.$$

The complex algebra corresponding to the reduct $(U, \wedge, \neg, \{L_B\}_{B\subseteq\mathcal{A}}, 0)$ of $\mathfrak{A}$ is obtained by extending the power set algebra over $PF(\mathfrak{A})$ with the operators $m_B^\mathfrak{A}$.

Now $\mathfrak{A}$ determines the lower approximation operators $\underline{Ind_B^{\mathfrak{A}_*}}$ induced by the CIS $\mathfrak{A}_*$ defined earlier. It also gives the complex algebra with operators $m_B^\mathfrak{A}$, as above. One can, in fact, show that for each $B \subseteq \mathcal{A}$, the operators $m_B^\mathfrak{A}$ and $\underline{Ind_B^{\mathfrak{A}_*}}$ are just the same. This result also leads to the representation theorem for the class of CIS-algebras.

Theorem 8. *Let* $\mathfrak{A}$ *be an abstract CIS-algebra. The map* $\Psi : U \to \mathcal{P}(PF(\mathfrak{A}))$ *defined for any* $x \in U$ *as*

$$\Psi(x) := \{\Gamma \in PF(\mathfrak{A}) : x \in \Gamma\},$$

is an embedding of $\mathfrak{A}$ *into* $(\mathfrak{A}_*)^*$.

4 Relationships

As observed in [29, 6], one can define to and fro transformations to show that pre-rough, regular double Stone, semi-simple Nelson and 3-valued Łukasiewicz algebras are all equivalent to each other.

It is not difficult to see that the defining axioms of pre-rough and 3-valued Łukasiewicz algebras (cf. Definitions 14 and 21 respectively), are deducible from each other.

The transformations involved for a passage to and from a pre-rough algebra $(A, \wedge, \vee, \neg, L, \Rightarrow, 0, 1)$ and a regular double Stone algebra (Definition 3) $(L, \vee, \wedge, ^*, ^+, 0, 1)$ are:

DP1. $a^+ := \neg La$,
DP2. $a^* := L(\neg a)$, and
PD1. $\neg a := (a \wedge a^+) \vee a^*$,
PD2. $La := a^{++}$.

For a semi-simple Nelson algebra (Definition 5) $\mathcal{N} := (A, \wedge, \vee, \neg, \sim, \rightarrow, 0, 1)$ and a pre-rough algebra $(A, \wedge, \vee, \neg, L, \Rightarrow, 0, 1)$, the transformations are:

NP1. $La = \neg \sim a$,
NP2. $a \Rightarrow b = \neg \sim (a \hookrightarrow b)$,
where $a \hookrightarrow b := (\sim a \wedge \sim \neg b) \vee (\neg \sim \neg a \vee b)$, and
PN1. $\sim a = \neg La$,
PN2. $a \rightarrow b = \neg La \vee b$.

It may be noted that an equivalent axiomatization of 3-valued Łukasiewicz algebras is given by the *Wajsberg algebras*.

Definition 29. *[11] A* Wajsberg algebra *is a structure* $(A, \rightarrow, \neg, 1)$ *such that*

1. $a \rightarrow (b \rightarrow a) = 1$,
2. $(a \rightarrow b) \rightarrow ((b \rightarrow c) \rightarrow (a \rightarrow c)) = 1$,
3. $((a \rightarrow \neg a) \rightarrow a) \rightarrow a = 1$,
4. $(\neg a \rightarrow \neg b) \rightarrow (b \rightarrow a) = 1$,
5. *If* $1 \rightarrow a = 1$ *then* $a = 1$,
6. *If* $a \rightarrow b = 1 = b \rightarrow a$ *then* $a = b$,

where $a, b, c \in A$.

Thus Wajsberg algebras also get related to the group of algebras being considered here. The transformations involved for a 3-valued Lukasiewicz algebra $(A, \leq, \sqcap, \sqcup, \neg, M, 0, 1)$ and a Wajsberg algebra $(A, \rightarrow, \neg, 1)$ are:

LW. $a \rightarrow b := (M\neg a \sqcup b) \sqcap (Mb \sqcup \neg a)$, and
WL1. $a \sqcup b := (a \rightarrow b) \rightarrow b$,
WL2. $a \sqcap b := \neg(\neg a \sqcup \neg b)$,
WL3. $Ma := \neg a \rightarrow a$,
WL4. $0 := \neg 1$.

A 3-valued Łukasiewicz algebra is cryptoisomorphic to an MV_3-algebra [26] in the sense of Birkhoff [9]. Thus all the preceding algebras are also cryptoisomorphic to MV_3-algebras as well.

Amongst algebras obtained from the operator-based approach, we observe the following relationships. Let U_B be the dual of the operator L_B in Definition 28, i.e. $U_B(x) := \neg L_B(\neg x)$. U_B and L_B are respectively closure and interior operators, and the reduct $\mathfrak{A} := (U, \wedge, \neg, \{L_B\}_{B \subseteq \mathcal{A}}, 0)$ is a topological Boolean algebra. $(U, \wedge, \neg, 0, \{U_B\}_{B \subseteq \mathcal{A}})$ satisfies all the conditions of an abstract knowledge approximation algebra (cf. Definition 25) except the following. In the latter case, the reduct $(U, \wedge, \neg, 0)$ is taken to be a complete atomic Boolean algebra, while we do not have that requirement. An abstract knowledge approximation algebra also needs to satisfy $U_{B \cup C}(x) = U_B(x) \wedge U_C(x)$, x being an atom, and this, in general, may not hold in an abstract CIS-algebra.

The difference between the signature of an abstract CIS-algebra of type $(\mathcal{A}, \mathcal{V})$, and that of a cylindric algebra of dimension $|\mathcal{A}|$ (cf. Definition 26) is the following. The cylindric algebra has unary and nullary operations corresponding to each element of $\mathcal{A}$ and $\mathcal{A} \times \mathcal{A}$ respectively. In the case of an abstract CIS-algebra, unary and nullary operations are indexed respectively over the sets $\mathcal{P}(\mathcal{A})$ and $\mathcal{A} \times \mathcal{V}$. Moreover, operators U_B of an abstract CIS-algebra satisfy (L_1)–(L_3), but may fail to satisfy (L_4). (L_5)–(L_7) do not make sense in the case of abstract CIS-algebras. However, the Boolean algebra $(U, \wedge, \neg, 0, U_B)$ with the operator U_B obtained from an abstract CIS-algebra, is a cylindric algebra of dimension 1.

5 Conclusions

We have presented algebraic structures that have been obtained so far, starting with set-based and operator-based definitions of Pawlak's rough sets. These have been abstracted, yielding known and new algebras, and the article summarizes results such as representation theorems for these algebras. The results indicate that some algebras known for years, get a new interpretation in terms of rough sets.

As mentioned in the Introduction, there has been work in the algebraic studies of generalized rough set theory also, and a number of interesting open questions remain. There are algebras of rough sets based on coverings, or defined on binary relations other than equivalences on the domain. Understandably, the picture gets more complicated there; in some cases the rough set structures do not even form lattices (cf. e.g. [20]). Algebras induced by an arbitrary number of approximation spaces on the same domain, called *multiple-source approximation system*, have been studied in [21]. Each approximation space in such a system represents the knowledge base of a source, hence the name. The algebraic investigation is open in case of finite *dynamic spaces* [30], essentially collections of a finite number of approximation spaces on the same domain. The interpretation in the case of these collections is that the knowledge base of a source is evolving over time [7].

Another direction of study is for classes of algebras induced by information systems that are not complete. Comer's work in [14] and the work in [22] are confined to complete information systems (CIS) only. But there are generalizations of CISs as well – we have *incomplete* and *non-deterministic* information systems, where some attribute value may not be known, or there may be multiple possibilities for the assignment of a value. Non-deterministic information systems induce a number of relations other than the standard indiscernibility relation on the domain. Considering these and abstracting [28], one gets different *information algebras*. However, as is the case with Comer's knowledge approximation algebras, these information algebras do not involve attribute and attribute-value pairs. So, as remarked earlier, these are unable to express the fact that approximations defined in the information systems are induced by attributes and their values. One of our current interests is to extend the notions and results of [22] to these generalized scenarios.

References

1. Banerjee, M.: A Categorial Approach to the Algebra and Logic of the Indiscernible. Ph.D. Thesis, University of Calcutta (1995)
2. Banerjee, M.: Rough sets and 3-valued Łukasiewicz logic. Fundamenta Informaticae **32**, 213–220 (1997)
3. Banerjee, M., Chakraborty, M. K.: A category for rough sets. Foundations of Computing and Decision Sciences **18**(3-4), 167–180 (1993)
4. Banerjee, M., Chakraborty, M. K.: Rough algebra. Bull. Polish Acad. Sc.(Math.) **41**(4), 293–297 (1993)
5. Banerjee, M., Chakraborty, M. K.: Rough sets through algebraic logic. Fundamenta Informaticae **28**(3-4), 211–221 (1996)
6. Banerjee, M., Chakraborty, M. K.: Algebras from rough sets. In: Pal, S. K., Polkowski, L., Skowron, A. (eds.) Rough-neuro Computing: Techniques for Computing with Words, pp. 157–184. Springer Verlag, Berlin (2004)
7. Banerjee, M. and Khan, M.A.: Rough set theory: a temporal logic view. In: Chakraborty, M.K. et al. (eds.) Studies in Logic Vol. 15, Proc. Logic, Navya-Nyaya & Applications: Homage to Bimal Krishna Matilal, Kolkata, 2007, pp. 1–20. College Publications, London (2008)
8. Banerjee, M. and Pal, K.: The variety of topological quasi-Boolean algebras. In: Trends in Logic III – Int. Conf. in Memoriam A. Mostowski, H. Rasiowa, C. Rauszer, Warsaw, Poland (2005). Extended abstract available at http://www.studialogica.org/.
9. Birkhoff, G.: Lattice Theory. AMS Colloquium 25, Providence, R.I. (1967)
10. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge Univ. Press (2001)
11. Boicescu, V., Filipoiu, A., Georgescu, G., Rudeano, S.: Łukasiewicz-Moisil Algebras. North Holland, Amsterdam (1991)
12. Bonikowski, Z.: A certain conception of the calculus of rough sets. Notre Dame J. of Formal Logic **33**, 412–421 (1992)
13. Cignoli, R.: Representation of Łukasiewicz and Post algebras by continuous functions. Colloq. Math. **24**(2), 127–138 (1972)

14. Comer, S.: An algebraic approach to the approximation of information. Fundamenta Informaticae **14**, 492–502 (1991)
15. Comer, S.: Perfect extensions of regular double Stone algebras. Algebra Universalis **34**, 96–109 (1995)
16. Epstein, G., Horn, A.: Chain based lattices. J. of Mathematics **55**(1), 65–84 (1974) Reprinted In: Rine, D. C. (ed.) Computer Science and Multiple-valued Logic. Theory and Applications, pp. 58–76. North Holland (1991)
17. Henkin, L., Monk, J. D., Tarski, A.: Cylindric Algebras, Part I. North-Holland, Amsterdam (1971)
18. Iturrioz, L.: Rough sets and three-valued structures. In: Orłowska, E. (ed.) Logic at Work: Essays Dedicated to the Memory of Helena Rasiowa, pp. 596–603. Physica-Verlag (1999)
19. Iwiński, T. B.: Algebraic approach to rough sets. Bull. Polish Acad. Sc. (Math.) **35**(9-10), 673–683 (1987)
20. Järvinen, J.: The ordered set of rough sets. In: Tsumoto, S. et al. (eds.) Proc. Rough Sets and Current Trends in Computing (RSCTC 2004), Sweden, 2004, LNAI 3066, pp. 49–58. Springer-Verlag (2004)
21. Khan, M.A. and Banerjee, M.: An algebraic semantics for the logic of multiple-source approximation systems. In: Sakai, H. et al. (eds.) Proc. Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC 2009), Delhi, India, 2009, LNCS 5908, pp. 69–76. Springer-Verlag (2009)
22. Khan, M.A. and Banerjee, M.: Information systems and rough set approximations: an algebraic approach. In: Kuznetsov, S.O. et al. (eds.) Proc. Pattern Recognition and Machine Intelligence (PReMI '11), Moscow, Russia, 2011, LNCS 6744, pp. 744–749. Springer-Verlag (2011).
23. Komorowski, J., Pawlak, Z., Polkowski, L. and Skowron, A.: Rough sets: A tutorial. In: Pal, S. K., Skowron, A. (eds.) Rough Fuzzy Hybridization: A New Trend in Decision–Making, pp. 3–98. Springer-Verlag, Singapore (1999)
24. Gehrke, M., Walker, E.: The structure of rough sets. Bull. Polish Acad. Sc. (Math.) **40**, 235–245 (1992)
25. Monteiro, A.: Construction des algebres de Łukasiewicz trivalentes dans les algebres de Boole monadiques. I. Math. Japonicae **12**, 1–23 (1967)
26. Mundici, D.: The C^*-algebras of three-valued logic. In: Ferro, R. et al. (eds.) Logic Colloquium '88, pp. 61–77. North Holland, Amsterdam (1989)
27. Obtułowicz, A.: Rough sets and Heyting algebra valued sets. Bull. Polish Acad. Sc. (Math.) **35**(9-10), 667–671 (1987)
28. Orłowska, E.: Introduction: what you always wanted to know about rough sets. In: Orłowska, E. (ed.) Incomplete Information: Rough Set Analysis, pp. 1–20. Physica-Verlag (1998)
29. Pagliani, P.: Rough set theory and logic-algebraic structures. In: Orłowska, E. (ed.) Incomplete Information: Rough Set Analysis, pp. 109–190. Physica-Verlag (1998)
30. Pagliani, P., Pretopologies and dynamic spaces, Fundamenta Informaticae, **59**(2-3), 221–239 (2004)
31. Pawlak, Z.: Rough sets. Int. J. Comp. Inf. Sci. **11**, 341–356 (1982)
32. Pawlak, Z.: Rough Sets. Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht (1991)
33. Pomykała, J. A.: Approximation, similarity and rough construction. Preprint, ILLC Prepublication Series CT–93–07, University of Amsterdam (1993)
34. Pomykała, J. and Pomykała, J. A.: The Stone algebra of rough sets. Bull. Polish Acad. Sc. (Math.) **36**, 495–508 (1988)

35. Rasiowa, H.: An Algebraic Approach to Non-classical Logics. North Holland, Amsterdam (1974)
36. Wasilewska, A., Vigneron, L.: Rough equality algebras. In: Wang, P. P. (ed.) Proc. Int. Workshop on Rough Sets and Soft Computing (JCIS'95), pp. 26–30 (1995)
37. Yao, Y.: Two views of the theory of rough sets in finite universes. Int. J. Approx. Reasoning **15**, 291–317 (1996)
38. Yao, Y.: Generalized rough set models. In: Polkowski, L., Skowron, A. (eds.) Rough Sets in Knowledge Discovery 1: Methodology and Applications, pp. 286–318. Physica-Verlag, Heidelberg (1998)

Mohua Banerjee
Department of Mathematics and Statistics,
Indian Institute of Technology, Kanpur 208016, India.
`mohua@iitk.ac.in`

Should Logic Be Normative?

Sanjukta Basu

1. Introduction

Should logic be normative? This question, like the question 'is this reasoning correct?' or, 'is this argument valid?' is a question which does not call for a matter-of-fact answer. This, rather, is a meta-normative question. In the wake of a number of non-classical logical systems during the past few decades, neither can we give a straight forward answer to the question 'is this argument valid?' by judging the argument's compliance with *the* rules for validity, nor can we pass off the claim that logic is normative without much ado. One reason for this is the plurality of logics in the current logical scenario, logics which oppose some laws/principles/rules of classical logic, and do not consider them to be valid in general. Pluralism in any field is an anathema to absolutism, its natural ally being relativism. Unless and until any of the proposed alternative systems of logic is capable to replace classical logic globally, or, the case for each such logic is shown to be untenable, relativism can not be avoided. Relativism undoubtedly would weaken the case for normativity of logic.

Quine's effort to naturalize epistemology[Quine, W.V.(1969)], and earlier, Hume's and other empiricists' naturalist explanation of morality have sought to shift the orientation of these studies from theoretical justification to scientific explanation of the phenomena under consideration. Likewise, for those who consider logical laws to be a posteriori and revisable in face of some empirical evidences, the question of normativity even in the context of logic is out of place. The above two possible cases of opposition to the absolute and a priori nature of logical laws represent two different aspects of non-normativity of logic. This paper aims to present these trends challenging classical logic, and to consider whether logic can be regarded as normative, and if so in what sense.

2. Traditional View of Logic as Normative

Traditionally it has been held that philosophy upholds three ideals – truth, beauty, and goodness.

Logic is supposed to be concerned with the ideal of truth. According to Aristotle, logic provides the organon for arriving at the most general truths about being-qua-being in his 'First Philosophy'. Thus logic is the study of those principles of reasoning/inference that would yield true conclusion from true premises; in other words, logic consists in the enlistment of truth-preserving, i.e. valid forms of arguments, and the fundamental laws of thought. That logic provides universal and necessary laws of thought, and Aristotle's syllogistics are the paradigm of correct reasoning was the general belief till the time of Kant. In his *Critique of Pure Reason* Kant remarks that all logic is a footnote on Aristotle even though many of these footnotes have today, in complexity, profundity, and range surpassed the whole of the original treatise. Classicists, including Kant hold that there is a strict distinction between the way we reason *actually* and the way we *ought* to reason as rational agents. Any attempt to conflate the fact-value dichotomy, and derive a value judgment from a factual statement is, it is said, to indulge in psychologism. Being a champion of anti-psycologism, Frege also considers logic to be concerned with the most general laws, which prescribe universally the way in which one ought to think if one is to think at all [Frege, G. (1893)].

Whether logic is conceived to be the science of the most general forms of valid inference, and the notion of consequence underlying those inferences or, logic is to be considered the system of logical truths – truths of the most general nature, a peculiar feature of logic in contrast to natural sciences like, physics is its normative character. The forms of inference or the laws that logic enunciate are not empirical generalizations, these are a priori. We would not enter here into a discussion of various views regarding the nature of logical laws. However, we must note that a Platonist like, Frege considers laws of logic e.g., the law of non-contradiction, the law of excluded middle, and the law of identity as the laws about the universal features of real entities, to be discovered by our logical intuition. In this sense, logic as embodying such laws is definitely an a priori and absolute science. Further, logic can be considered normative in the sense that it prescribes those principles of inference as correct, which accord with the universal and a priori laws of reality. According to Kant however, laws of logic are the a priori conditions of our understanding and thought. But on this view it is not plausible how humans could think at all except in accordance with the laws of logic. In that case it would be meaningless to say that we ought to obey these laws.

Why the rules of inference that classical logic prescribes are the rules of correct reasoning? This is because these rules preserve truth. And, though we sometimes believe what is false, our beliefs can not be considered as knowledge unless we believe truly. In this regard, Hartry Field [Field, H. (2009)] observes that our views about entailment or logical consequence constrain our views about how we ought to reason or, about the proper interrelations among our beliefs. The classical notion of validity as preservation of truth is actually a deterrent against what one should not believe as true; that is, it warns us, as if, against believing not-B when we know or, at least believe that A implies B, and that A. If we do not follow this prohibition, we will believe not only something false, but more significantly, we would end up nurturing a system of beliefs, which may be inconsistent. The notion of validity as necessary truth-preservation acts as a safeguard against allowing any inferential move in a formal system that would enable one to derive contradiction (A and not-A) in that system. However, this classical notion of validity as necessary truth-preservation, which is captured model theoretically as preservation of truth in all models, is not an unrestricted notion of validity as the Curry paradox shows [Field, H. (2009)].

3. Pluralism in Logic

But, even when the classical notion of validity is restricted to the bounds of a given theory and not allowed to step into the purview of its meta-theory, and to apply to inferences expressible in the meta-language, some of the classical rules of inference / laws are allegedly not validated by this notion of validity. Examples abound. Various systems of many-valued logic, fuzzy logic, intuitionist logic, quantum logic, relevance logic, para-consistent logic particularly, dialeathic logic, and so on – an enumeration of all these logics that have opposed classical logic in some way or other is perplexing enough to doubt whether logic is normative, and more importantly, to doubt the feasibility of this very question. The main suspects among the classical laws / rules are the law of non-contradiction (challenged by various systems of many-valued logic, and systems of para-consistent logic), the rule for double negation (to infer A from not not-A, denied by intuitionist logic), the law of distributivity (contested by quantum logic), the law of excluded middle (opposed by various systems of many-valued logic, and also, rejected by intuitionist logic), even the classical notion of validity and consequence that permits any proposition to follow logically from a contradiction have been decried and rejected (relevance logic, para-consistent logic).

Now, rejection of any of the classical logical laws / rules also precipitates change of meaning of the logical constants involved essentially in those laws / rules. In this context, Quine's opinion regarding the predicament of the deviant logicians is to be noted. Quine says that when a deviant logician denies, say, the law of non-contradiction by considering some conjunction of the form 'p and not-p' as true, and checks derivation of any sentence in the proposed system from such a conjunction by adjusting the existing rules of derivation, he thinks he is talking about negation, 'not', but surely 'not' has ceased to be recognizable as negation. "When he tries to deny the doctrine he only changes the subject" [Quine, W.V.(1970), p.81].

Even so, we must say, as Quine himself admits that the issue between classical logic and non-classical logic is not just verbal. In repudiating the law of excluded middle – 'p or not-p' – the deviant logician is indeed giving up classical negation, or perhaps alternation, or both; and he may have his reason. The more important issue in this context is whether this reason for change of meaning is such as would recommend a thorough change or, would recommend restricted and localized change in the meanings of the logical constants. In the former case there would be a genuine rivalry between classical logic and non-classical logic. In the latter case there would be no genuine opposition between classical logic and non-classical logic or, between any two non-classical logics. Each logic including, classical logic having its own field of enquiry would co-exist side by side peacefully. There would be no incompatibility in effect between them. Change of meaning in such cases would turn out to be a case of equivocation or relabelling. For example, the law of distributivity is no longer held to be valid in quantum mechanics, and a quantum logician should use different symbols for conjunction and disjunction from those that are used in classical logic. However, outside the scope of quantum mechanics quantum logician has no qualms to admit the law of distributivity, and there he uses 'and' and 'or' in their old classical meanings. A thorough-going pragmatism would be a natural consequence of this kind of relativistic pluralism in logic. For those who subscribe to this view, the question of normativity of logic is beside the point.

But according to Dummett, whether non-classical logic should replace classical logic is an issue pertaining to the theory of meaning. Thus, the issue between classical logic and intuitionist logic relates to "the correct model for the meanings which we confer upon our mathematical statements. A model of meaning, in this sense, is a model of understanding, that is, of what it is to know the meaning of a statement" [Dummett, M. (1976), p.288]. This knowledge of meaning, according to Dummett, has to relate to "the means available to us for knowing the truth of statements of the relevant class: in the quantum- logical case, in terms of measurement of physical qualities; in the intuitionistic case, in terms of proofs of mathematical propositions" [Dummett, M. (1976), pp.288 – 289].Thus, on Dummett's view, the question which logic is the correct logic of reasoning is a meaningful question. An answer to this question depends on a choice of the correct theory of meaning. And which theory is the correct theory of meaning is a question which is "irreducibly philosophical in character".

4. Naturalizing Logic

The most prominent effort in naturalizing logic, in questioning the demarcation between statements which are true by virtue of meanings of words alone (analytic), and statements which are true solely by virtue of matter of fact (synthetic) is also made by Quine [Quine, W.V. (1951), (1969)]. The scathing attack against the fundamentality of the distinction between analytic and synthetic statements in 'Two Dogmas of Empiricism' [Quine, W. V. (1951)], together with the theory of indeterminacy of meaning and meaning holism prompted Quine to regard even logic not to be immune to any revision in the theory in which logic lies at the core, due to recalcitrant empirical evidence. Putnam too joined Quine in a league against a priority of logic [Putnam, H. (1968)]. Indeed, how much this extremism in Quine is theoretical is to be judged from his later views. Thus even conceding to the possibility of revision of logical laws on empirical grounds, Quine [Quine, W. V. (1978), p.81] holds: "If sheer logic is not conclusive, what is? What higher tribunal could abrogate the logic of truth functions or of quantification?" Then, with a pragmatic note he maintains that since a revision in physical theory, however extensive it may be, will always be less disruptive of our total theory than a revision in classical logic, the former will always be preferred to the latter.

In recent times, works in artificial intelligence and belief revision, results obtained from experiments like, Wason's card selection task have made some philosophers and cognitive scientists to realize that everyday reasoning is much more sensitive to the context of particular statements than formal logic is. They hold that if a content is believable then it tends to be believed whether logic or any other normative standards dictate us to do so or not. It can be said, however, that in such a case the statement is believed on some ground other than logic. It is believed not as a conclusion of an invalid argument, the premises of which have nothing to do with the truth or the degree of belief of that statement. The fact that majority of us, while engaged in reasoning do not follow classical logical rules as the data collected from experiments on Wason's tasks seek to establish, indicate on the contrary that logical rules are prescriptive in nature.

Logic is of course not experimental or even theoretical psychology; it approaches human reasoning with a different purpose of its own. But a logical theory can not be totally disjoint from actual reasoning, also. We often have to draw conclusion from incomplete inadequate information, which calls for making provisions in the tools of reasoning for withdrawing previous conclusions, and even making conclusions that might contradict already existing premises. This led researchers in the area of artificial intelligence and belief revision to abandon the condition of monotonicity of consequence. Perhaps a redefinition of the term 'logic' should be on the agenda of those who think logic should be responsive to context.

5. Conclusion

The debate concerning the role of empirical evidence in changing classical logical rules, regarding the fact-value dichotomy, and also, regarding absolutism vs. relativism will go on. Philosophers will continue debating whether an archipelago of distinct purpose-oriented systems of logic is to be recognized, or logicians should search for an all-purpose unitary logic. Meanwhile, we must admit that all rational practices, all rational discourses, all theory building have to abide by some regulative principles which are the common minimum desiderata of mutual understanding and sharing of views, of the very possibility and plausibility of any debate between the contesting parties.

References

1. Dummett, M. (1976): 'Is Logic Empirical?' in M.Dummett, *Truth and Other Enigmas*, London: Duckworth, 1978.
2. Field, H. (2009): 'Pluralism in Logic', *The Review of Symbolic Logic*, Vol. 2, No. 2, June, 2009, pp. 342 – 359.
3. Frege, G. (1893): *Grundgesetze der Arithmetic* translated by Montgomery Furth as *Basic Laws of Arithmetic*, Berkeley: University of California Press, 1967.
4. Kant, Immanuel (1974): *Logic: Immanuel Kant*, translated with an introduction by R.S. Hartman and W. Schwartz, New York: Dover Publications, Inc.
5. Putnam, H. (1968): 'Is Logic Empirical?' in R.S.Cohen and M. Wartofsky (Eds.) *Boston Studies for Philosophy of Science*, vol.v, pp. 216 – 241.
6. Quine, W.V. (1950): *Methods of Logic,* New York: Holt, Rinehart and Winston Inc.
7. Quine, W.V. (1951): 'Two Dogmas of Empiricism', *Philosophical Review*, 60/1 January; reprinted in W.D. Hart (Ed.), *Philosophy of Mathematics*, Oxford: Oxford University Press, 1976, pp.31-51.
8. Quine, W.V.(1969): 'Epistemology Naturalized', in W.V.Quine, Ontological Relativity and Other Essays, New York: Columbia University Press, pp. 69-90.
9. Quine, W.V. (1970): Philosophy of Logic, USA: Prentice-Hall Inc., Englewood Cliffs; Indian reprint, New Delhi: Prentice-Hall of India Pvt. Ltd., 1978.

Internet Resources

1. Johan van Benthem (2007): 'Logic and reasoning: do the facts matter?', http://staff.science.uva.nl/~johan.
2. Field, H. (2009): 'What is the Normative Role of Logic?', Proceedings of the Aristotelian Society, Supplementary, 83(1), pp. 251-268; also available online, http://philpapers.org/rec/FIEWIT

Sanjukta Basu
Department of Philosophy,
Rabindrabharati University, Kolkata, India.
juktab@yahoo.co.in

Rough Sets and Vague Concepts

Jan Bazan, Andrzej Skowron, Jarosław Stepaniuk
and Roman Swiniarski

Abstract. In this paper, we summarize and extend our previous results on relationships between rough sets and vague concepts (see, e.g., [87, 5, 104, 63, 92]). In particular, we consider imperfect specifications of vague concepts (e.g., by examples) and changes of knowledge about vague concepts. In these cases, the rough set based methods are allowing us to induce temporary approximations of vague concepts. These approximations may change with changes of knowledge about the approximated concepts. Approximation spaces used for concept approximation have been initially defined on samples of objects (decision tables) representing partial information about concepts. Such approximation spaces defined on samples are next inductively extended on the whole object universe. This makes it possible to define the concept approximation on extensions of samples. We discuss the role of inductive extensions of approximation spaces in searching for concept approximation. However, searching for relevant inductive extensions of approximation spaces defined on samples is infeasible for compound concepts. We outline an approach making this searching feasible by using a concept ontology specified by domain knowledge and its approximation. We also extend this approach to a framework for adaptive approximation of vague concepts by agents interacting with environments. This paper realizes a step toward approximate reasoning in multiagent systems (MAS), intelligent systems, and complex dynamic systems (CAS).

1 Introduction

In this paper we discuss the rough set approach to vague concept approximation.

There is a long debate in philosophy on vague concepts [36]. Nowadays, computer scientists are also interested in vague (imprecise) concepts, e.g, many intelligent systems should satisfy some constraints specified by vague concepts. Hence, the problem of vague concept approximation as well as preserving vague dependencies especially in dynamically changing environments is important for such systems. Lotfi Zadeh [117] introduced a very successful approach to vagueness. In this approach, sets are defined by partial membership in contrast to crisp membership used in the classical definition of a set. Rough set theory [59] expresses vagueness not by means of membership but by employing the boundary region of a set. If the boundary region of a set is empty it means that a particular set is crisp, otherwise the set is rough (inexact). The non-empty boundary region of the set means that our knowledge about the set is not sufficient to define the set precisely. In this paper some consequences on understanding of vague concepts caused by inductive extensions of approximation spaces and adaptive concept learning are presented. A discussion on vagueness in the context of fuzzy sets and rough sets can be found in [74].

Initially, the approximation spaces were introduced for decision tables (samples of objects). The assumption was made that the partial information about objects is given by values of attributes and on the basis of such information about objects the approximations of subsets of objects form the universe restricted to sample have been defined [59]. Starting, at least, from the early 90s, many researchers have been using the rough set approach for constructing classification algorithms (classifiers) defined over extensions of samples. This is based on the assumption that available information about concepts is partial. In recent years, there have been attempts based on approximation spaces and operations on approximation spaces for developing an approach to approximation of concepts over the extensions of samples (see, e.g., [89, 93, 94, 104]). In this paper, we follow this approach and we show that the basic operations related to approximation of concepts on extension of samples are inductive extensions of approximation spaces. For illustration of the approach we use approximation spaces defined in [88]. Among the basic components of approximation spaces are neighborhoods of objects defined by the available information about objects and rough inclusion functions between sets of objects. Observe that searching for relevant (for approximation of concepts) extensions of approximation spaces requires tuning many more parameters than in the case of approximation of concepts on samples. The important conclusion from our considerations is that the inductive extensions defining classification algorithms (classifiers) are defined by arguments "for" and "against" of concepts. Each argument is defined by a tuple consisting of a degree of inclusion of objects into a pattern and a degree of inclusion of the pattern into the concepts. Patterns in the case of rule-based classifiers can be interpreted as the left hand sides of decision rules. The arguments are discovered from available data and can be treated as the basic information granules used in the concept approximation process. For any new object, it is possible to check the satisfiability of arguments and select arguments satisfied to a satisfactory degree. Such selected arguments are fused by conflict resolution strategies for obtaining the classification decision. Searching for relevant approximation spaces in the case of approximations over extensions of samples requires discovery of many parameters and patterns including selection of relevant attributes defining information about objects, discovery of relevant patterns for approximated concepts, selection of measures (similarity or closeness) of objects into discovered patters for concepts, structure and parameters of conflict resolution strategy. This causes that in the case of more compound concepts the searching process becomes infeasible (see, e.g., [14, 113]). We propose to use as hints in the searching for relevant approximation spaces for compound concepts an additional domain knowledge making it possible to approximate such concepts. This additional knowledge is represented by a concept ontology [7–9, 50–52, 85, 86, 89, 90, 105] including concepts expressed in natural language and some dependencies between them. We assume that the ontology of concept has a hierarchical structure. Moreover, we assume that for each concept from ontology there is given a labelled set of examples of objects. The labels show the membership for objects relative to the approximated concepts. The aim is to discover the relevant conditional attributes for concepts on different levels of a hierarchy. Such attributes can be constructed using the so-called production rules, productions, and approximate reasoning schemes (AR schemes, for short) discovered from data (see, e.g. [7–9, 50–52, 85, 86, 89, 90, 105])). The searching for relevant arguments "for" and "against" for more compound concepts can be simplified because the searching can be organized along the derivations over the ontology using the domain knowledge.

Notice, that the searching process for relevant approximation spaces is driven by some selected quality measures. While in some learning problems such measures can be selected in a relatively easy way and remain unchanged during learning in other learning processes they can be only approximated on the basis of a partial information about such measures, e.g., received as the result of interaction with the environment. This case concerns, e.g., adaptive learning. We discuss the process of searching for relevant approximation spaces in different tasks of adaptive learning [2, 19, 29, 32, 46, 48, 42, 106]. In particular, we present illustrative examples of adaptation of observation to the agent's scheme, incremental learning, reinforcement learning, and adaptive planning. Our discussion is presented in the framework of multiagent systems (MAS). The main conclusion is that the approximation of concepts in adaptive learning requires much more advanced methods, which, in particular, will make it possible to approximate the quality measures together with approximation of concepts. We suggest that this approach can be also based on approximation of ontology. In adaptive learning, the approximation of concepts is constructed gradually and the temporary approximations are changing dynamically in the learning process in which we are trying to achieve the approximation of the relevant quality. This, in particular, causes, e.g., boundary conditions to change dynamically during the learning process in which we are attempting to find the relevant approximation of the boundary regions of approximated vague concepts. This is consistent with the requirement of the higher order vagueness [36] stating that the borderline cases of vague concepts are not crisp sets. In Conclusions we point out some consequences of this fact for further research on the rough set logic.

This paper is an extension and continuation of several papers (see, e.g., [7–9, 50–52, 87, 85, 86, 89, 90, 93, 104, 5, 96, 92]) on approximation spaces and vague concept approximation processes. In particular, we discuss here a problem of adaptive learning of concept approximation. In this case, we are also searching for relevant approximation of the quality approximation measure. In a given step of the learning process, we have only a partial information about such a measure. On the basis of such information we construct its approximation and we use it for inducing approximation spaces relevant for concept approximation. However, in the next stages of the learning process, it may happen that after receiving new information form the environment, it is necessary to reconstruct the approximation of the quality measure and in this way we obtain a new "driving force" in searching for relevant approximation spaces during the learning process.

This paper is organized as follows. Section 2 presents an introductory discussion on sets and vagueness. In Section 3, we discuss inductive extensions of approximation spaces. We emphasize the role of discovery of special patterns and the so called arguments in inductive extensions. In Section 4, the role of approximation spaces in hierarchical learning is presented. Section 4, outlines and approach based on approximation spaces during adaptive learning. In Section 6 (Conclusions), we summarize the discussion presented in the paper and we present some further research directions based on approximation spaces to approximate reasoning in multiagent systems and complex adaptive systems.

2 Sets and Vague Concepts

In this section, we put forward some general remarks on the concept of a set and the place of rough sets in set theory.

The concept of a set is fundamental for the whole of mathematics. Modern set theory was formulated by George Cantor [16]. Bertrand Russell discovered that the intuitive notion of a set proposed by Cantor leads to antinomies [79]. Two kinds of remedy for this problem have been proposed: axiomatization of Cantorian set theory and alternative set theories.

Another issue discussed in connection with the notion of a set or a concept is vagueness (see, e.g., [11, 24, 36, 37, 77, 80]). Mathematics requires that all mathematical notions (including set) must be exact (Gottlob Frege [26]). However, philosophers (see, e.g., [36, 37, 74, 77]) and recently computer scientists (see, e.g., [59, 117]) have become interested in vague concepts.

2.1 Sets

The notion of a set is one of, if not the most, basic concepts in mathematics. All mathematical structures refer to it.

The definition of this notion and the creation of set theory are due to German mathematician Georg Cantor (1845-1918) [16], who laid the foundations of contemporary set theory over 100 years ago.

The birth of set theory can be traced back to his proof in 1873 of the uncountability of real line (i.e., the set of all real numbers is not countable) [15]. It was Bernhard Bolzano (1781-1848) who coined the term *Menge* ("set"), which Cantor used to refer to objects in his theory. According to Cantor, a set is a collection of any objects, which can be considered as a whole according to some law. As one can see, the notion of set is very intuitive and simple.

Mathematical objects such as relations, functions, numbers, are examples of sets. In fact, set theory is needed in mathematics to provide rigor.

The notion of a set is not only fundamental for the whole of mathematics but it also plays an important role in natural language. We often speak about sets (collections) of various objects of interest such as collection of books, paintings and people.

The intuitive meaning of a set according to some dictionaries is the following: "A number of things of the same kind that belong or are used together."

Webster's Dictionary

"Number of things of the same kind, that belong together because they are similar or complementary to each other."

The Oxford English Dictionary

Thus a set is a collection of things which are somehow related to each other but the nature of this relationship is not specified in these definitions.

In fact, these definitions are due to the original definition given by Cantor.

2.2 Antinomies

In 1903, the renowned English philosopher Bertrand Russell (1872-1970) observed [79] that the intuitive notion of a set given by Cantor leads to logical *antinomies* (contradictions), i.e., Cantor set theory is contradictory (there are other kinds of antinomies, which are outside the scope of this paper). A logical antinomy (for simplicity, we refer to "antinomy" in the rest of this paper) arises whenever correct logical reasoning leads to a contradiction, i.e., to propositions A and *non-*A, which is not allowed in logic.

As an example let us discuss briefly the so-called Russell's antinomy. Consider the set X containing all the sets Y, which are not the elements of themselves. If we assume that X is its own element then X, by definition, cannot be its own element; while if we assume that X is not its own element then, according to the definition of the set X, it must be its own element. Thus while applying each assumption we obtain contradiction.

Antinomies show that a set cannot be a collection of arbitrary elements, as was stipulated by Cantor.

One could think that antinomies are ingenuous logical play, but it is not so. They question the essence of logical reasoning. That is why there have been attempts to "repair" Cantor's theory for over 100 years or to substitute another set theory for it but the results so far are still not satisfactory. So, is all mathematics based on doubtful foundations?

As a remedy for this defect several axiomatizations of set theory have been proposed (e.g., Zermelo and Fraenkel, 1904 [33]).

Instead of improvements of Cantors' set theory by its axiomatization, some mathematicians proposed escape from classical set theory by creating a completely new idea of a set, which would free the theory from antinomies [41, 114, 1].

No doubt the most interesting proposal was given by Polish logician Stanisław Leśniewski, who introduced the relation of "being a part" instead of the membership relation between elements and sets employed in classical set theory. In his set theory called mereology, *being a part* is a fundamental relation [41]. Mereology is a significant part of recent studies on the foundations of mathematics (see, e.g., [12, 17, 25, 112, 39, 82]), artificial intelligence [98], cognitive science [97], natural language [23], and research in rough set theory (see, e.g., [70, 71, 67]).

The problem of finding an alternative to classical set theory has failed to be solved until now.

The deficiency of sets, mentioned above, has rather philosophical than practical meaning, since sets used practically in mathematics are free from the above discussed faults. Antinomies are associated with very "peculiar" sets constructed in logic but not found in sets used in "everyday" mathematics. That is why we can use mathematics safely.

2.3 Vagueness

Another issue discussed in connection with the notion of a set is vagueness. Mathematics requires that all mathematical notions (including set) must be exact, otherwise precise reasoning would be impossible. However, philosophers [36, 37, 74, 78] and recently computer scientists (see, e.g., [3, 53, 54, 45, 18, 61, 62, 87, 5, 22, 63, 101, 13, 69, 96, 92]) as well as other researchers have become interested in *vague* (imprecise) concepts.

In classical set theory a set is uniquely determined by its elements. In other words, this means that every element must be uniquely classified as belonging to the set or not. That is to say the notion of a set is a *crisp* (precise) one. For example, the set of odd numbers is crisp because every number is either odd or even.

In contrast to odd numbers, the notion of a beautiful painting is vague, because we are unable to classify uniquely all paintings into two classes: beautiful and not beautiful. Some paintings cannot be decided whether they are beautiful or not and thus they remain in the doubtful area. Thus, *beauty* is not a precise but a vague concept.

Almost all concepts in natural language are vague. Therefore, common sense reasoning based on natural language must be based on vague concepts and not on classical logic. Interesting discussion of this issue can be found in [74].

The idea of vagueness can be traced back to the ancient Greek philosopher Eubulides of Megara (ca. 400BC) who first formulated so called "sorites" (heap) and "falakros" (bald man) paradoxes (see, e.g., [36, 37]). The bald man paradox goes as follows: suppose a man has 100,000 hairs on his head. Removing one hair from his head surely cannot make him bald. Repeating this step we arrive at the conclusion the a man without any hair is not bald. Similar reasoning can be applied to a hip of stones.

Vagueness is usually associated with the boundary region approach (i.e., existence of objects which cannot be uniquely classified relative to a set or its complement) which was first formulated in 1893 by the father of modern logic, German logician, Gottlob Frege (1848-1925) (see [26]).

According to Frege the concept must have a sharp boundary. To the concept without a sharp boundary there would correspond an area that would not have any sharp boundary–line all around. It means that mathematics must use crisp, not vague concepts, otherwise it would be impossible to reason precisely.

Summing up, vagueness is

– not allowed in mathematics;
– interesting for philosophy;
– nettlesome problem for natural language, cognitive science, artificial intelligence, machine learning, and computer science.

Discussion on vague (imprecise) concepts includes the following issues [36]:

1. The presence of borderline cases.
2. Boundary regions of vague concepts are not crisp.
3. Vague concepts are susceptible to sorites paradoxes.

The existence of boundary cases are represented in rough sets by boundary regions which are defined as crisp sets. However, boundary regions are defined relative to a given set of attributes and in a more general setting to approximation space used for approximation. Hence, one can obtain many different approximations of vague concepts with many boundary regions. Note that these are not all possible boundary regions. Next, many different methods for aggregating the boundary regions can be used. In selection of these aggregating methods again some subjective or inductive assumptions are used. In the consequence, due to imperfect knowledge about the approximated concept, there is no possibility to define using classical set theory the absolute boundary region consisting all borderline cases for a given concept. In applications, the approximation of vague concepts is induced using some optimization criteria, e.g., based on the minimum length principle. The obtained boundary region can be treated as a temporary approximation of the absolute boundary region. This view may be changed, if our knowledge about the approximated concept was changed. For example, it is usually assumed that the understanding (approximation) of vague concepts (their semantics is determined by the satisfiability relation) depends on the agent's knowledge, which is often changing. Hence, the approximation of vague concepts by an agent should also be considered changing with time (this is known as the concept drift). This problem is discussed in more detail in this paper.

One can use rough sets for explanations of the existence of borderline instances discussed in the literature (see, e.g., [110, 18]). If x is and object and C is a concept than using the rough set approach the following situations may happen:

1. x is either C or not C but it is not known which case is true. This situation may happen if in the decision table there is an object with unknown decision assigned by, e.g., domain expert. Note that any further conditional attributes can not change this situation.
2. x is actually neither C nor $non - P$. This may happen if one considers only certain knowledge about C and then in may happen that the union of the

lower approximation of C and the lower approximation of the complement of C is not equal to the whole universe of objects. This situation may be also happen for induced classifiers. For example, in the case of rule based classifiers, weights of rules matched by x which are voting for C and against C may be not discernible to a satisfactory degree [50].
3. x is partially C, and partially $non - C$. By introducing the rough membership function one can obtain membership degrees different from 0 and 1 for objects in the boundary region.
4. Depending on the context (perspective), x is sometimes C, and sometimes $non-C$. This situation may happen when we deal with inconsistent decision tables [63]. Following some context, sometimes one can add some new (conditional) attributes and with these new attributes the decision table becomes consistent.

3 Approximation Spaces and their Inductive Extensions

In [59], any approximation space is defined as a pair (U, R), where U is a universe of objects and $R \subseteq U \times U$ is an indiscernibility relation defined by an attribute set.

The lower approximation, the upper approximation and the boundary region are defined as crisp sets. It means that the higher order vagueness condition is not satisfied [36]. We will return to this issue in Section 5.

We use the definition of approximation space introduced in [88]. Any approximation space is a tuple $AS = (U, I, \nu)$, where U is the universe of objects, I is an uncertainty function, and ν is a measure of inclusion called the inclusion function, generalized in rough mereology to the rough inclusion [88, 94].

3.1 Uncertainty Function

In [91], the uncertainty function I defines for every object x from a given sample U of objects, a set of objects with descriptions similar to x. The set $I(x)$ is called the neighborhood of x. In our next papers (see, e.g. [95, 4, 92]) more general uncertainty functions are considered for structural objects. The structures of objects are induced in hierarchical modeling. Let us recall some basic concepts.

Let

$$P_\omega(U^*) = \bigcup_{i \geq 1} P^i(U^*),$$

where $P^1(U^*) = P(U^*)$ and $P^{i+1}(U^*) = P(P^i(U^*)))$ for $i \geq 1$. For example, if $card(U^*) = 2$ and $U^* = \{x_1, x_2\}$, then we obtain $P^1(U^*) = \{\emptyset, \{x_1\}, \{x_2\}, \{x_1, x_2\}\}$, $P^2(U^*) = \{\emptyset, \{\emptyset\}, \{\{x_1\}\}, \{\{x_2\}\}, \{\{x_1, x_2\}\}, \{\emptyset, \{x_1\}\}, \{\emptyset, \{x_2\}\}, \{\emptyset, \{x_1, x_2\}\}, \ldots\}$, ... If $card(U^*) = n$, where n is a positive natural number, then $card(P^1(U^*)) = 2^n$ and $card(P^{n+1}(U^*)) = 2^{card(P^n(U^*))}$, for $n \geq 1$. For example, $card(P^3(U^*)) = 2^{2^{2^n}}$. Hence, we see that the levels of the powerset hierarchy are very rich and full automatic searching for relevant sets (structures) on such levels is not feasible. However, this approach allows us to present in a uniform way foundations for modeling of granular computing aimed at inducing compound granules from different levels of the powerset hierarchy relevant for solving the target task, e.g., approximation of complex concepts. For applications, it is necessary to restrict searching for relevant granules in relevant fragments of the powerset hierarchy. These fragments are defined by some sets of formulas. Discovery of such sets often is the big challenge for many problems.

Now, we consider uncertainty functions of the form $I : U^* \longrightarrow P_\omega(U^*)$. The values of uncertainty functions are called granular neighborhoods. These granular neighborhoods are defined by the so called granular formulas. The values of such uncertainty functions are not necessarily from $P(U^*)$ but from $P_\omega(U^*)$. Figure 1 presents an illustrative example of the uncertainty function with values in $P^2(U^*)$ rather than in $P(U^*)$. The generalization of neighborhoods discussed here are also motivated by the necessity of modeling or discovery of complex structural objects in solving problems of pattern recognition, machine learning, or data mining. These structural objects (granules) can be defined as sets on higher levels of the powerset hierarchy. Among examples of such granules are indiscernibility or similarity classes of patterns or relational structures discovered in images, clusters of time widows, indiscernibility or similarity classes of sequences of time windows representing processes, behavioral graphs (for more details see, e.g., [95, 4]).

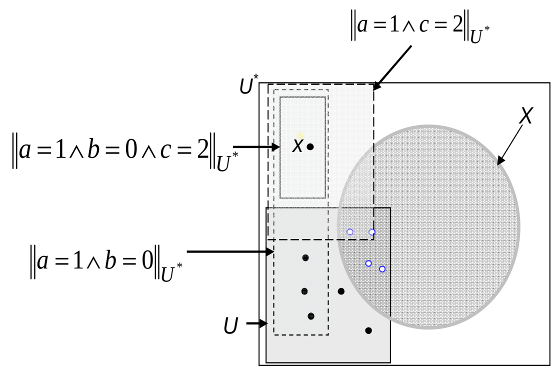


Fig. 1. Uncertainty function $I : U^* \rightarrow P^2(U^*)$. The neighborhood of $x \in U^* \setminus U$, where $Inf_A(x) = \{(a, 1), (b, 0), (c, 2)\}$, does not contain training cases from U. The generalizations of this neighborhood described by formulas $\|a = 1 \wedge c = 2\|_{U^*}$ and $\|a = 1 \wedge b = 0\|_{U^*}$ have non empty intersections with U.

If $X \in P_\omega(U^*)$ and $U \subseteq U^*$, then by $X \upharpoonright U$ we denote the set defined as follows (i) if $X \in P(U^*)$ then $X \upharpoonright U = X \cap U$ and (ii) for any $i \geq 1$ if $X \in P^{i+1}(U^*)$ then $X \upharpoonright U = \{Y \upharpoonright U : Y \in X\}$. For example, if $U = \{x_1\}$,

$U^* = \{x_1, x_2\}$ and $X = \{\{x_2\}, \{x_1, x_2\}\}$ $(X \in P^2(U^*))$, then $X \upharpoonright U = \{Y \upharpoonright U : Y \in X\} = \{Y \cap U : Y \in X\} = \{\emptyset, \{x_1\}\}$.

In the above example, we discussed uncertainty functions assigning granules from $P^2(U^*)$ to objects from U^* . We assume that objects from U and U^* are of atomic type with the type defined by attributes from a given information system $IS = (U, A)$ over U. For example, the type of objects from U may be identified with $\times_{a \in A} V_a$. In a number of papers (see, e.g., [92]), we present methods of defining information systems with objects of higher order type (structural objects). Such objects can be from different levels of $P_\omega(U^*)$, i.e., they can belong to $P^k(U^*)$, for $k > 2$. Note that neighborhoods (e.g., indiscernibility classes) over such objects are sets of objects of higher order type from $P_\omega(U^*)$.

3.2 Rough Inclusion Function

The second component of any approximation space is the rough inclusion function [91].

One can consider general constraints which the rough inclusion functions should satisfy. In this section, we present only some examples of rough inclusion functions.

The rough inclusion function $\nu : P(U) \times P(U) \to [0, 1]$ defines the degree of inclusion of X in Y, where $X, Y \subseteq U$ and U is a finite sample of objects.

In the simplest case the standard rough inclusion function ν_{SRI} can be defined by (see, e.g., [91], [63]):

$$\nu_{SRI}(X, Y) = \begin{cases} \frac{card(X \cap Y)}{card(X)}, & \text{if } X \neq \emptyset \\ 1, & \text{if } X = \emptyset. \end{cases} \tag{1}$$

Some illustrative examples are given in Table 1.

Table 1. Illustration of Standard Rough Inclusion Function

X	Y	$\nu_{SRI}(X, Y)$
$\{x_1, x_3, x_7, x_8\}$	$\{x_2, x_4, x_5, x_6, x_9\}$	0
$\{x_1, x_3, x_7, x_8\}$	$\{x_1, x_2, x_4, x_5, x_6, x_9\}$	0.25
$\{x_1, x_3, x_7, x_8\}$	$\{x_1, x_2, x_3, x_7, x_8\}$	1

It is important to note that an inclusion measure expressed in terms of the confidence measure, widely used in data mining, was considered by Łukasiewicz [43] a long time ago in studies on assigning fractional truth values to logical formulas.

For a definition of the inclusion function for more general granules, e.g., partitions of objects, one can use a measure based on positive region [63], entropy [27] or rough entropy [56, 44]. Inclusion measures for more general granules were also investigated [84, 10]. However, more work in this direction should be done, especially on inclusion of granules with complex structures, in particular for granular neighborhoods.

In this section, we consider the problem of approximation of concepts over a universe U^*, i.e., subsets of U^*. We assume that the concepts are perceived only through some subsets of U^*, called samples. This is a typical situation in machine learning, pattern recognition, or data mining [27]. In this section we explain the rough set approach to induction of concept approximations. The approach is based on inductive extension of approximation spaces.

Now we will discuss in more detail the approach presented in [93, 94]. Let $U \subseteq U^*$ be a finite sample and let $C_U = C \cap U$ for any concept $C \subseteq U^*$. Let $AS = (U, I, \nu)$ be an approximation space over the sample U. The problem we consider is how to extend the approximations of C_U defined by AS to approximation of C over U^*.

In Figure 2, formula α is representing the description of the induced classification algorithm (classifier) for the concept C and its semantics of $\|\alpha\|_U$ on sample U and semantics $\|\alpha\|_{U^*}$ on U^*. The aim is to induce a formula α with semantics consistent not only with given examples on sample U (positive from $C \cap U$ and negative from $(U^* \setminus C) \cap U$) but also such that its extension $\|\alpha\|_{U^*}$ on U^* is as close as possible to the concept C on U^* .

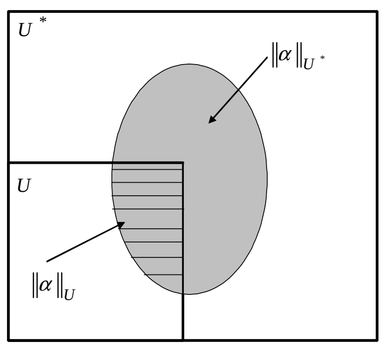


Fig. 2. Two semantics of formula α: $\|\alpha\|_U$ and α: $\|\alpha\|_{U^*}$ on the sample U and its extension, respectively U^*

We show that the problem can be described as searching for an extension $AS^* = (U^*, I^*, \nu^*)$ of the approximation space AS relevant for approximation of C. This requires showing how to induce values of the extended inclusion function to relevant subsets of U^* that are suitable for the approximation of C. Observe that for the approximation of C, it is enough to induce the necessary values of the inclusion function ν^* without knowing the exact value of $I^*(x) \subseteq U^*$ for $x \in U^*$.

We consider an example for rule-based classifiers[1]. However, the analogous considerations for k-NN classifiers, feed-forward neural networks, and hierarchical classifiers [27] show that their construction is based on the inductive inclusion extension [87, 94].

Usually, neighborhoods of objects in approximation spaces are defined by some formulas called patterns. Let us consider an example. Let AS^* be a given approximation space over U^* and let us consider a language L of patterns, where x denotes an object from U^*. In the case of rule-based classifiers, patterns are defined by feature value vectors. More precisely, in this case any pattern $pat(x)$ is defined by a formula $\bigwedge\{(a, a(x)) : a \in A \text{ and } v_a \in V_a\}$, where A is a given set of condition attributes [59]. An object $u \in U^*$ is satisfying $\bigwedge\{(a, a(x)) : a \in A \text{ and } v \in V_a\}$ if $a(u) = a(x)$ for any $a \in A$, i.e., if and only if x, u are A-indiscernible [59]. The set of objects satisfying $pat(x)$ in U^*, i.e., the semantics of $pat(x)$ in U^*, is denoted by $\|pat(x)\|_{U^*}$. Hence, $\|pat(x)\|_{U^*} = [x]_A$ where $[x]_A$ is the A-indiscernibility class of $x \in U^*$ [59]. By $\|pat(x)\|_U$ we denote the restriction of $\|pat(x)\|$ to $U \subseteq U^*$, i.e., the set $\|pat(x)\| \cap U$. In the considered case, we assume that any neighborhood $I(x) \subseteq U$ in AS is is expressible by a pattern $pat(x)$. It means that $I(x) = \|pat(x)\|_U \subseteq U$, where $\|pat(x)\|_U$ denotes the meaning of $pat(x)$ restricted to the sample U.

We assume that for any object $x \in U^*$, only partial information about x resulting, e.g., from a sensor measurement represented by a pattern $pat(x) \in L$ with semantics $\|pat(x)\|_{U^*} \subseteq U^*$ defining the neighborhood of x in U^* is available. Moreover, only partial information such as $\|pat(x)\|_U$ is available about this set. In particular, relationships between information granules over U^*, e.g., $\|pat(x)\|_{U^*}$ and $\|pat(y)\|_{U^*}$, for different $x, y \in U^*$, are known only to a degree estimated by using relationships between the restrictions of these sets to the sample U, i.e., between sets $\|pat(x)\|_{U^*} \cap U$ and $\|pat(y)\|_{U^*} \cap U$.

The set $\{pat(x) : x \in U\}$ of patterns (defined by the whole set of attributes A from an approximation space AS) is usually not relevant for approximation of the concept $C \subseteq U^*$. Such patterns can be too specific or not general enough, and can directly be applied only to a very limited number of new sample elements. For example, if for a new object $x \in U^* \setminus U$ the set $\|pat(x)\|_U$ can be disjoint with all sets $\|pat(y)\|_U$ for $y \in U$.

However, by using some generalization strategies, one can induce from patterns belonging to $\{pat(x) : x \in U\}$ some new patterns that are relevant for approximation of concepts over U^*.

Usually, first we define a new set PAT of patterns, which are candidates for relevant approximation of a given concept C. A typical example of the set of such patterns used in the case of rule based classifiers can be defined by dropping some descriptors from patterns constructed over the whole set of attributes, i.e., $\{\bigwedge\{(a, a(x)) : a \in B \text{ and } v_a \in V_a\} : B \subseteq A \text{ and } x \in U\}$. Among such patterns we search for the left hand sides of decision rules.

[1] For simplicity of reasoning we consider only binary classifiers, i.e. classifiers with two decision classes. One can easily extend the approach to the case of classifiers with more decision classes.

The set $PATTERNS(AS, L, C)$ can be selected from PAT using some quality measures evaluated on meanings (semantics) of patterns from this set restricted to the sample U. Often such measures are based on the numbers of examples from the concept C_U and its complement that support (satisfy) a given pattern. For example, if the confidence coefficient

$$\frac{card(\|pat\|_U \cap C_U)}{card(\|pat(x)\|_U)}, \tag{2}$$

where $pat \in PAT$, is at least equal to a given threshold and the support

$$\frac{card(\|pat\|_U \cap C_U)}{card(U}, \tag{3}$$

is also at least equal to a given threshold than we select pat as a member of $PATTERNS(AS, L, C)$.

Next, on the basis of some properties of sets definable by patterns from $PATTERNS(AS, L, C)$ over U, we induce approximate values of the inclusion function $\nu^*(X, C)$ on subsets of $X \subseteq U^*$ definable by any such pattern and the concept C. For example, we assume that the value of the confidence coefficient is not changing significantly if we move from U to U^*, i.e.,

$$\frac{card(\|pat\|_U \cap C_U)}{card(\|pat(x)\|_U)} \approx \frac{card(\|pat\|_{U^*} \cap C)}{card(\|pat(x)\|_{U^*})}, \tag{4}$$

Next, we induce the value of ν^* on pairs (X, Y) where $X \subseteq U^*$ is definable by a pattern from $\{pat(x) : x \in U^*\}$ and $Y \subseteq U^*$ is definable by a pattern from $PATTERNS(AS, L, C)$. For example, if $pat(x) = \bigwedge\{(a, a(x)) : a \in A \text{ and } v_a \in V_a\}$ and pat is obtained from $pat(x)$ by dropping some conjuncts then $\nu^*(\|pat(x)\|_{U^*}, \|pat\|)_{U^*}) = 1$ because $\|pat\|_{U^*} \subseteq \|pat(x)\|_{U^*}$. In a more general case, one can estimate the degree of inclusion of $\|pat(x)\|_{U^*}$ into $\|pat\|_{U^*}$ using some similarity degrees defined between formulas from $PATTERNS(AS, L, C)$ and PAT. For example, one can assume that the values of attributes on x which occur in pat are not necessarily the same but similar. Certainly, such a similarity should be also defined or learned from data.

Finally, for any object $x \in U^* \setminus U$ we induce the degree $\nu^*(\|pat(x)\|_{U^*}, C)$ applying a conflict resolution strategy $Conflict\_res$ (e.g, a voting strategy) to two families of tuples:

$$\{(\nu^*(\|pat(x)\|_{U^*}, \|pat\|_{U^*}), pat, \nu^*(\|pat\|_{U^*}, C)) : pat \in PATTERNS(AS, L, C)\}. \tag{5}$$

Let us observe that conflicts can occur due to inductive reasoning in estimation of values of ν^*. For some $x \in U^*$ and $pat, pat' \in PATTERNS(AS, L, C)$ the values $\nu^*(\|pat(x)\|_{U^*}, \|pat\|_{U^*})$, $\nu^*(\|pat(x)\|_{U^*}, \|pat'\|_{U^*})$ can be both large (i.e., close to 1) and at the same time the value $\nu^*(\|pat\|_{U^*}, C)$ can be small (i.e., close to 0) and the value of $\nu^*(\|pat'\|_{U^*}, C)$ can be large.

Values of the inclusion function for the remaining subsets of U^* can be chosen in any way – they do not have any impact on the approximations of C. Moreover, observe that for the approximation of C we do not need to know the exact values of uncertainty function I^* – it is enough to induce the values of the inclusion function ν^*. The defined extension ν^* of ν to some subsets of U^* makes it possible to define an approximation of the concept C in a new approximation space AS^*.

To reduce the number of conditions from (5) one can use the so called arguments "for" and "against" discussed, e.g., in [90].

Any C-argument, where $C \subseteq U^*$ is a concept is a triple

$$(\epsilon, pat, \epsilon') \tag{6}$$

where $\epsilon, \epsilon' \in [0, 1]$ are degrees and pat is a pattern from $PATTERNS(AS, L, C)$.

The argument $arg = (\varepsilon, pat, \varepsilon')$ is satisfied by a given object $x \in U^*$, in symbols $x \models_C arg$, if and only if the following conditions are satisfied:

$$\begin{aligned} &\nu^*(\|pat(x)\|_{U^*}, \|pat\|_{U^*}) \geq \varepsilon; \\ &\nu^*(\|pat\|_{U^*}, C) \geq \varepsilon'. \end{aligned} \tag{7}$$

The idea of C-arguments is illustrated in Figure 3.

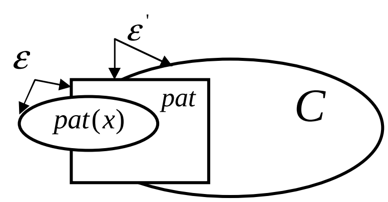


Fig. 3. C-argument

Instead of all conditions from (5) only some arguments "for" and "against" C are selected and the conflict resolution strategy is based on them. For any new object the strategy resolves conflicts between selected arguments "for" and "against" C which are satisfied by the object.

The very simple strategy for selection of arguments is the following one. The C-argument $arg = (\varepsilon, pat, \varepsilon')$ is called the argument "for" C if $\varepsilon \geq t$ and $\varepsilon' \geq t'$, where $t, t' > 0.5$ are given thresholds. The argument arg is "against" C, if this argument is the argument for the complement of C, i.e., for $U^* \setminus C$. However, in general this may be not relevant method and the relevant arguments should be selected on the basis of more advanced quality measures. They can take into account , e.g., the support of patterns in arguments (see Section 5.1), their coverage, independence from other arguments, or relevance in searching for arguments used for approximation of more compound concepts in hierarchical learning.

One can define the lower approximation and upper approximation of the concept $C \subseteq U^*$ in the approximation space AS^* by

$$LOW(AS^*, C) = \{x \in U^* : \nu^*(I^*(x), C) = 1\}, \tag{8}$$
$$UPP(AS^*, C) = \{x \in U^* : \nu^*(I^*(x), C) > 0\}.$$

From the definition we have:

$$LOW(AS^*, C) \cap U \subseteq C \cap U \subseteq UPP(AS^*, C) \cap U. \tag{9}$$

However, in general the following equalities do not hold:

$$LOW(AS, C \cap U) = LOW(AS^*, C) \cap U, \tag{10}$$
$$UPP(AS, C \cap U) = UPP(AS^*, C) \cap U.$$

One can check that in the case of standard rough inclusion [89] we have:

$$LOW(AS, C \cap U) \supseteq LOW(AS^*, C) \cap U, \tag{11}$$
$$UPP(AS, C \cap U) \subseteq UPP(AS^*, C) \cap U.$$

Following the minimal length principle [75, 76, 100] some parameters of the induced approximation spaces are tuned to obtain a proper balance between the description length of the classifier and its consistency degree. The consistency degree on a given sample U of data can be represented by degrees to which the sets defined in equalities (10) are close. The description length is measured by description complexity of the classifier representation. Among parameters which are tuned are attribute sets used in the classifier construction, degrees of inclusion of patterns defined by objects to the left hand sides of decision rules, degrees of inclusion of patterns representing the left hand sides of decision rules in the decision classes, the specificity or support of these patterns, parameters of the conflict resolution strategy (e.g., set of arguments and parameters of arguments).

We can summarize our considerations in this section as follows. The inductive extensions of approximation spaces are basic operations on approximation spaces in searching for relevant approximation spaces for concept approximation. The approximation of concepts over U^* is based on searching for relevant approximation spaces AS^* in the set of approximation spaces defined by inductive extensions of a given approximation space AS. For any object $x \in U^* \setminus U$, the value $\nu^*(I^*(x), C)$ of the induced inclusion function ν^* is defined by conflict resolution strategy from collected arguments *for* classifying x to C and from collected arguments *against* classifying x to C.

4 Approximation Spaces in Hierarchical Learning

The methodology for approximation spaces extension presented in Section 3 is widely used for construction of rule based classifiers. However, this methodology cannot be directly used for concepts that are compound because of problems with inducing of the relevant set $PATTERNS(AS, L, C)$ of patterns. For such compound concepts, hierarchical learning methods have been developed (see, e.g., [6–9, 50–52, 85, 86, 89, 90, 105]).

We assume that domain knowledge is available about concepts. There is given a hierarchy of concepts and dependencies between them creating the so-called *concept ontology*. Only partial information is available about concepts in the hierarchy. For concepts from the lowest level of hierarchy, decision tables with condition attributes representing sensory measurements are given. Classifiers for these concepts are induced (constructed) from such decision tables. Assuming that classifiers have been induced for concepts from some level l of the hierarchy, we are aiming at inducing classifiers for concepts on the next $l+1$ level of the hierarchy. It is assumed that for concepts on higher levels there are given samples of objects with information about their membership values relative to the concepts. The relevant patterns for approximation of concepts from the $l+1$ level are discovered using (i) these decision tables, (ii) information about dependencies linking concepts from the level $l+1$ with concepts from the level l, and (iii) patterns discovered for approximation of concepts from the level l of the hierarchy. Such patterns define condition attributes (e.g., by the characteristic functions of patterns) in decision tables. Next, from such approximation of concepts are induced. In this way, also, the neighborhoods for objects on the level $l+1$ are defined. Observe also that the structure of objects on the higher level $l+1$ is defined by means of their parts from the level l. In this section, for simplicity of reasoning, we assume that on each level the same objects are considered. To this end, we also assume that rough inclusion functions from approximation spaces are standard rough inclusion functions [89].

Now we outline a method of construction of patterns used for approximation of concepts from a given level of concept hierarchy by patterns used for approximation of concepts belonging to the lower level of the hierarchy. This approach has been elaborated in a number of papers cited above, in particular in [90]. Assume that a concept C belongs to a level $l+1$ of the hierarchy. We outline the idea of searching for sets $PATTERNS(AS, L, C)$ of patterns for a concept C, where AS is an approximation space discovered for approximation of the concept C and L is a language in which discovered patterns are expressed.

To illustrate this idea, let us consider and example of a dependency for a concept C from domain knwoledge:

$$\text{if } C_1 \text{ and } C_2 \text{ then } C, \tag{12}$$

where C_1, C_2, C are vague concepts. Analogously, let us consider a dependency for the complement of C:

$$\text{if } C_1' \text{ and } C_2' \text{ then } \neg C. \tag{13}$$

In general, we should consider a set with many dependencies with different concepts on the right hand sides of dependencies (creating, e.g., a partition of the universe) and in the process of generating arguments "for" and "against" a selected concept C are involved other vague dependencies from the given set. Let us recall that such a set of concepts and dependencies between them is specified in a given domain knowledge.

To approximate the target concept C, relevant patterns for C and $\neg C$ should be derived. The main idea is presented in Figure 4 and Figure 5. We assume that for any considered concept and for each pattern selected for this concept a degree of its inclusion into the concept can be estimated.

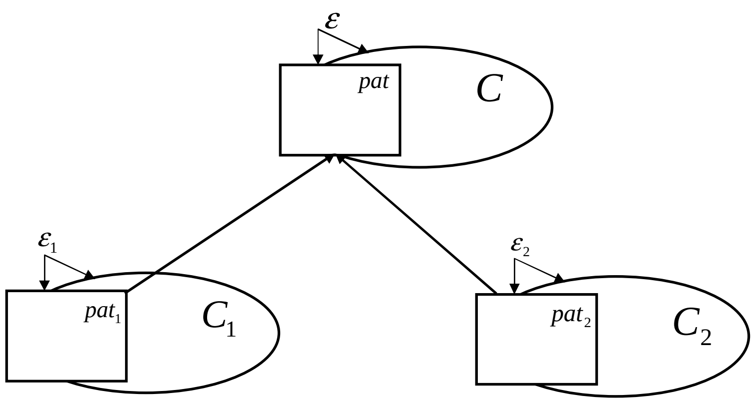


Fig. 4. An illustration of pattern construction

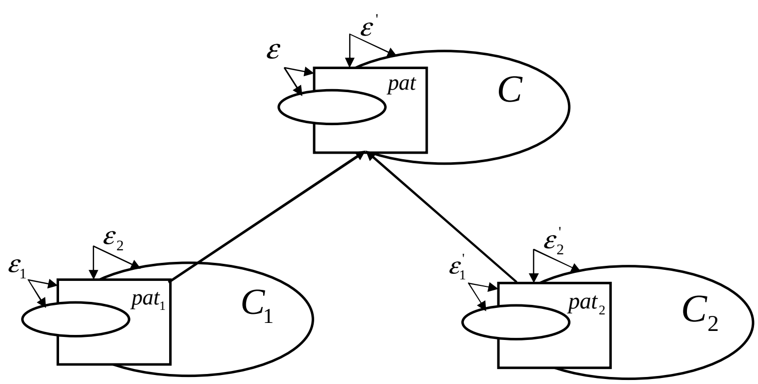


Fig. 5. An illustration of production rule

In Figure 4 it is shown that for patterns pat_1, pat_2 (e.g., left hand sides of decision rules in case of a rule based classifiers) for (or against) C_1 and C_2 and their inclusion degrees ϵ_1 and ϵ_2 into C_1 and C_2, respectively, it is constructed a pattern *pat* for (or against) C together with estimation of its inclusion degree ϵ to the concept C.

Figure 5 represents a construction of the target argument $(\epsilon, pat, \epsilon')$ for C from arguments $(\epsilon_1, pat_1, \epsilon'_1)$ and $(\epsilon_2, pat_2, \epsilon'_2)$ for C_1 and C_2, respectively. Such a construction

$$\text{if } (\epsilon_1, pat_1, \epsilon'_1) \text{ and } (\epsilon_2, pat_2, \epsilon'_2) \text{ then } (\epsilon, pat, \epsilon') \tag{14}$$

is called a *production rule* for the dependency (12). Such a production rule is true at a given object x if and only if the following implication holds:

$$\text{if } x \models_{C_1} (\epsilon_1, pat_1, \epsilon'_1) \text{ and } x \models_{C_2} (\epsilon_2, pat_2, \epsilon'_2) \text{ then } x \models_C (\epsilon, pat, \epsilon'). \tag{15}$$

Certainly, it is necessary to search for production rules of the high quality (with respect to some measures) making possible to construct "strong" arguments in the conclusion of the production from "strong" arguments in the premisses of the production rule. The quality of arguments is defined by means of relevant degrees of inclusion in these arguments and properties of patterns (such as support or description length).

The quality of arguments for concepts from the level $l+1$ can be estimated on the basis properties of arguments for the concepts from the level l from which these arguments have been constructed. In this estimation are used decision tables delivered by domain experts. Such decision tables consist of objects with decision values equal to the membership degrees of objects relative to the concept or to its complement. In searching for productions of high quality, we use operations called *constrained sums* (see, e.g., [99]). Using these operations there are performed joins of information systems representing patterns appearing in arguments from the premise of production. The join is parameterized by constraints helping by tuning these parameters to filter the relevant objects from composition of patterns used for constructing a pattern for the concept C on the level $l+1$ for the argument in the conclusion of the production rule. Moreover, the production rules may be composed into the so called approximation reasoning schemes (AR schemes, for short). This makes it possible to generate patterns for approximation of concepts on the higher level of the hierarchy (see, e.g., [6–9, 50–52, 86, 90]). In this way one can induce gradually for any concept C in the hierarchy a relevant set of arguments (based on the relevant set of patterns $PATTERNS(AS, L, C)$ of patterns; see Section 4) for approximation of C.

We have recognized that for a given concept $C \subseteq U^*$ and any object $x \in U^*$, instead of crisp decision about the relationship of $I^*(x)$ and C, we can gather some arguments *for* and *against* it only. Next, it is necessary to induce from such arguments the value $\nu^*(I(x), C)$ using some strategies making it possible to resolve conflicts between those arguments [27, 89]. Usually some general principles are used such as the minimal length principle [27] in searching for algorithms computing an extension $\nu^*(I(x), C)$. However, often the approximated concept over $U^* \setminus U$ is too compound to be induced directly from $\nu(I(x), C)$. This is the reason that the existing learning methods are not satisfactory for inducing high quality concept approximations in case of complex concepts [113]. There have been several attempts trying to omit this drawback. In this section we have discussed the approach based on hierarchical (layered) learning [105].

There are some other issues which should be discussed in approximation of compound vague concepts. Among them are issues related to adaptive learning and construction or reconstruction of approximation spaces in interaction with environments. In the following section, we consider an agent learning some concepts. This agent is learning the concepts in interaction with the environments. Different types of interaction are defining different types of adaptive learning processes. In particular one can distinguish incremental learning [30, 47, 109, 116], reinforcement learning [21, 31, 35, 64, 73, 108], competitive or cooperative learning [32]. There are several issues, important for adaptive learning that should be mentioned. For example, the compound target concept which we attempt to learn can gradually change over time and this concept drift is a natural extension for incremental learning systems toward adaptive systems. In adaptive learning it is important not only what we learn but also how we learn, how we measure changes in a distributed environment and induce from them adaptive changes of constructed concept approximations. The adaptive learning for autonomous systems became a challenge for machine learning, robotics, complex systems, and multiagent systems. It is becoming also a very attractive research area for the rough set approach. Some of these issues will be discussed in the following section.

5 Approximation Spaces in Adaptive Learning

There are different interpretations of the terms adaptive learning and adaptive systems (see, e.g., [2, 19, 29, 32, 46, 48, 42, 106]). We mean by adaptive a system that learns to change with its environment. Our understanding is closest to the spirit of what appears in [19, 29]. In complex adaptive systems (CAS), agents scan their environment and develop a schema for action. Such a schema defines interactions with agents surrounding it together with information and resources flow externally [19]. In this section, we concentrate only on some aspects of adaptive learning. The other issues of adaptive learning in MAS and CAS will be discussed elsewhere.

In particular, we would like to discuss the role of approximation spaces in adaptive learning.

In this paper, we consider the following exemplary situation. There is an agent *ag* interacting with another agent called the environment (ENV). Interactions are represented by actions [28, 111] performed by agents. These actions are changing values of some sensory attributes of agents. The agent *ag* is equipped with ontology of vague concepts consisting of vague concepts and dependencies between them.

There are three main tasks of the agent *ag*: (i) adaptation of observation to the agent's scheme, (ii) adaptive learning of the approximations of vague concepts, and (iii) preserving constraints (e.g., expressed by dependencies between concepts).

Through adaptation of observation to the agent's scheme agent becomes more robust and can handle more variability [19].

Approximation of vague concepts by the agent ag requires development of searching methods for relevant approximation spaces which create the basis for approximation of concepts. Observe that the approximations of vague concepts are dynamically changing in adaptive learning when new knowledge about approximated concept is obtained by the agent ag. In particular, from this it follows that the boundary regions of approximated concepts are dynamically changing in adaptive learning. For each approximated concept we obtain a sequence of boundary regions rather than a single crisp boundary region. By generating this sequence we are attempting to approximate the set of borderline cases of a given vague concept. Hence, if the concept approximation problem is considered in adaptive framework the higher order postulate for vague concepts is satisfied (i.e., the set of borderline cases of any vague concept can not be crisp) [36, 87, 93].

The third task of the agent ag requires learning of a planning strategy. This is a strategy for predicting plans (i.e., sequences of actions) on the basis of observed changes in the satisfiability of the observed concepts from ontology. By executing plans the actual state of the system is transformed to a state satisfying the constraints. Changes in the environments can cause that the executed plans should be reconstructed dynamically by relevant adaptive strategies. In our example actions performed by the agent ag are adjusting values of sensory attributes which are controllable by ag.

Before we will discuss the mentioned above tasks in more detail we would like to add some comments on interaction between agents.

The interactions among agents belong to the most important ingredients of computations realized by multiagent systems [42]. In particular, adaptive learning agents are interact, in particular, with their environments. In this section, we will continue our discussion on adaptive learning by agents interacting with environment. Some illustrative examples of interactions which have influence on the learning process are presented.

Let us consider two agents ag and ENV representing the agent learning some concepts and the environment, respectively. By $ag\_s(t)$ and $ENV\_s(t)$ we denote (information about) the state of agents ag and ENV at the time t, respectively. Such an information can be represented, e.g., by a vector of attribute values A_{ag} and A_{ENV}, respectively [94]. The agent ag is computing the next state $ag_s(t+1)$ using his own transition relation $\longrightarrow_{ag}$ applied to the result of interaction of $ag_s(t)$ and $ENV_s(t)$. The result of such an interaction we denote by $ag_s(t) \oplus_{ENV} ENV_s(t)$ where $\oplus_{ENV}$ is an operation of interaction of ENV on the state of ag. Hence, the following condition holds:

$$ag_s(t) \oplus_{ENV} ENV_s(t) \longrightarrow_{ag} ag_s(t+1). \tag{16}$$

Analogously, we obtain the following transition for environment states:

$$ag_s(t) \oplus_{ag} ENV_s(t) \longrightarrow_{ENV} ENV_s(t+1). \tag{17}$$

In our examples, we will concentrate on two examples of interactions. In the first example related to incremental learning (see, e.g., [47, 109, 30, 116]), we assume that $ag_s(t) \oplus_{ENV} ENV_s(t)$ is obtained by extending of $ag_s(t)$ by a new information about some new sample of objects labelled by decisions. The structure of $ag_s(t)$ is much more compound than in non-incremental learning. This will be discussed in one of the following section together with some aspects of adaptation in incremental learning. These aspects are related to searching for relevant approximation spaces. In the discussed case, we also assume that $ag_s(t) \oplus_{ag} ENV_s(t)$, i.e., there is no interaction of the agentag on the environment. In our second example, the agent ag can change the state of ENV by performing some actions or plans which change the state of the environment.

5.1 Adaptation of observation to the agent's scheme

In this section, we present two illustrative examples of adaptation of observation to the agent's scheme. In the consequence of such an adaptation, the agent's scheme becomes more robust relative to observations.

In the first example, we consider instead of patterns $pat(x)$ (see Section 3) mode general patterns which are obtained by granulation of such patterns using a similarity relation τ. Assuming that the object description $pat(x)$ is defined by $\bigwedge\{(a, a(x)) : a \in A \text{ and } v_a \in V_a\}$ one can define such a similarity τ on description of objects, e.g., by a composition of similarity relations on attribute value sets (see, e.g., [40, 49, 88])[2]. Then instead of patterns $pat(x)$ we obtain patterns $pat_\tau(x)$ with the semantics defined by $\|pat_\tau(x)\|_{U^*} = \{y \in U^* : pat(x)\tau pat(y)\}$. The definition of satisfiability of arguments (7) changes as follows

$$\begin{aligned}&\nu^*(\|pat_\tau(x)\|_{U^*}, \|pat\|_{U^*}) \geq \varepsilon;\\&\nu^*(\|pat\|_{U^*}, C) \geq \varepsilon'.\end{aligned} \tag{18}$$

Observe, that $\|pat_\tau(x)\|_{U^*}$ is usually supported by many more objects than $\|pat(x)\|_{U^*}$. Hence, if it is possible to tune the parameters of τ in such a way that the first condition in (18) is satisfied for sufficiently large ε than the obtained argument is much more robust than the previous one, i.e., it is satisfied by much more objects than the previous one $pat(x)$ and at the same time the requirement related to the degrees of inclusion is preserved.

Our second example concerns construction of more robust production rules and productions (sets of production rules corresponding to the same dependency between vague concepts) (see Figure 6). Patterns in such productions represent different layers of vague concepts and are determined by the linguistic values of membership such as *small, medium, high* (see, e.g., [9]). These more general patterns are constructed using information granulation [94]. Let us consider a

[2] Note, that the similarity relation τ has usually many parameters which should be tuned in searching for relevant similarity relations.

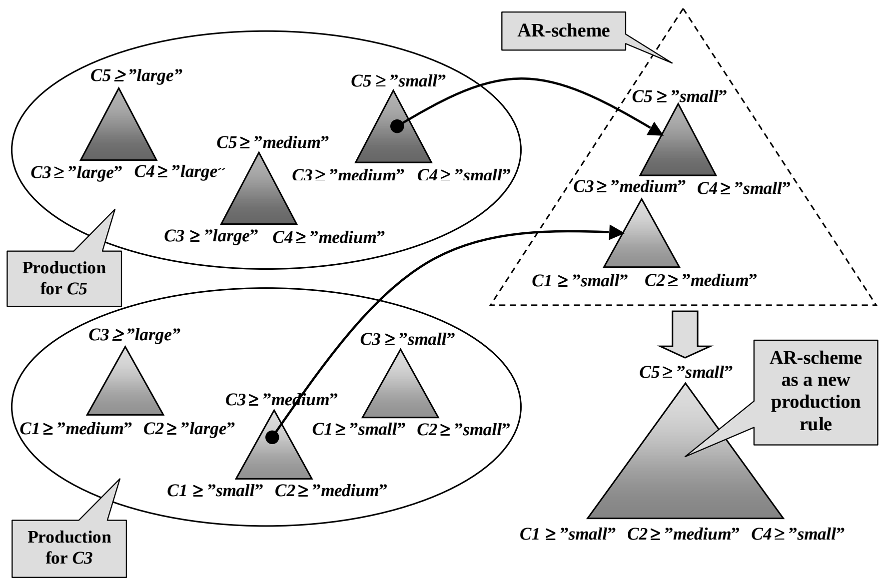


Fig. 6. An illustration of production and AR scheme

simple example of information granulation. Observe that the definition of the satisfiability of arguments given by (7) is not unique. One can consider the decision table (U, A, d), where A is a given set of condition attributes [59] and the decision d is the characteristic function of the set $Y_\varepsilon(pat) = \{y \in U : \nu(\|pat(y)\|_U, \|pat\|_U)) \geq \varepsilon\}$. From this decision table can be induced the classifier $Class(pat)$ for the concept $Y^*_\varepsilon(pat) = \{y \in U^* : \nu^*(\|pat(y)\|_{U^*}, \|pat\|_{U^*})) \geq \varepsilon\}$. Any object $x \in U^*$ is satisfying the C-argument (6) if and only if the following condition is satisfied:

$$\nu^*(Y^*_\varepsilon(pat), C) \geq \varepsilon'. \tag{19}$$

The satisfiability of (19) is estimated by checking if the following condition holds on the sample U:

$$\nu(Y_\varepsilon(pat), C \cap U) \geq \varepsilon'. \tag{20}$$

We select only the arguments $(\varepsilon, pat, \varepsilon')$ with the maximal ε' satisfying (20) for given ε and pat.

Assume that $0 = \varepsilon_0 < \ldots < \varepsilon_{i-1} < \varepsilon_i < \ldots < \varepsilon_n = 1$. For any $i = 1, \ldots, n$ we granulate a family of sets

$$\{Y^*_\varepsilon(pat) : pat \in PATTERNS(AS, L, C) \text{ and } \nu^*(Y^*_\varepsilon(pat), C) \in [\varepsilon_{i-1}, \varepsilon_i)\} \tag{21}$$

into one set $Y^*_\varepsilon(\varepsilon_{i-1}, \varepsilon_i)$. Each set $Y^*_\varepsilon(\varepsilon_{i-1}, \varepsilon_i)$ is defined by an induced classifier $Class_\varepsilon(\varepsilon_{i-1}, \varepsilon_i)$. The classifiers are induced, in an analogous way as before, by constructing a decision table over a sample $U \subseteq U^*$. In this way we obtain a family of classifiers $\{Class_\varepsilon(\varepsilon_{i-1}, \varepsilon_i)\}_{i=1,\ldots,n}$.

The sequence $0 = \varepsilon_0 < \ldots < \varepsilon_{i-1} < \varepsilon_i < \ldots < \varepsilon_n = 1$ should be discovered in such a way that the classifiers $Class_\varepsilon(\varepsilon_{i-1}, \varepsilon_i)$ correspond to different layers of the concept C with linguistic values of membership. One of the method in searching for such sequence can be based on analysis of a histogram. This histogram represents a function $f(I)$ where $I \in \mathcal{J}$, $\mathcal{J}$ is a given uniform partition of the interval $[0,1]$, and $f(I)$ is the number of patterns from $\{Y_\varepsilon^*(pat) : pat \in PATTERNS(AS, L, C)\}$ with the inclusion degree into C from $I \subseteq [0,1]$.

5.2 Adaptation and incremental learning

In this section, we outline a searching process for relevant approximation spaces in incremental learning. Let us consider an example of incremental concept approximation scheme Sch (see Figure 7). By $Inf(C)$ and $Inf'(C)$ we denote

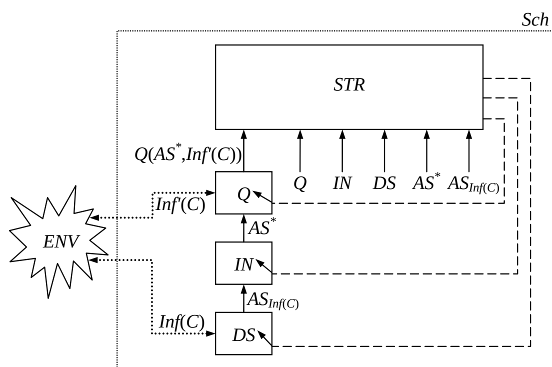


Fig. 7. An example of incremental concept approximation scheme

a partial information about the approximated concept (e.g., decision table for C or training sample) in different moments t and $t+1$ of time, respectively[3]. ENV denotes an environment, DS is an operation constructing an approximation space $AS_{Inf(C)}$ from a given sample $Inf(C)$, i.e., a decision table. IN is an inductive extension operation (see Section 3) transforming the approximation space $AS_{Inf(C)}$ into an approximation space AS^* used for approximation of the concept C; Q denotes an operation returning the quality of the induced approximation space AS^* on a new sample $Inf'(C)$, i.e., an extension of the decision table $Inf(C)$. For example, the value $Q(AS^*, Inf'(C))$ can be taken as a ratio

[3] For simplicity, in Figure 7 we do not present time constraints.

of the number of objects from $Inf'(C)$ that are classified correctly (relative to the decision values from $Inf'(C)$) by the classification algorithm (classifier) for C defined by AS^* (see Section 3) to the number of all objects in $Inf'(C)$.

The double-ended arrows leading into and out of ENV illustrate an interaction between agent ag and the environment ENV. In the case of a simple incremental learning strategy only samples of C are communicated by ENV to ag. More compound interactions between ag and ENV will be discuss later. They can be related to reaction from ENV on predicted by ag decisions (actions, plans) (see, e.g., award and penalty policies in reinforcement strategies [21, 31, 35, 64, 73, 108]).

STR is a strategy that adaptively changes the approximation of C by modifying the quality measure Q, the operation of inductive extension IN, and the operation DS of constructing the approximation space $AS_{Inf(C)}$ from the sample $Inf(C)$. Dotted lines outgoing from the box labelled by STR in Figure 7 are illustrating that the strategy STR after receiving the actual values of input its parameters is changing them (e.g., in the next moment of time). To make Figure 7 more readable the dotted lines are pointing to only one occurrence of each parameter of STR but we assume that its occurrences on the input for STR are modified too.

In the simple incremental learning strategy, the quality measure is fixed. The aim of the strategy STR is to optimize the value of Q in the learning process. This means that in the learning process we would like to reach as soon as possible an approximation space which will guarantee the quality of classification measured by Q to be almost optimal. Still, we do not know how to control by STR this optimization. For example, should this strategy be more like the annealing strategy [38], then it is possible to perform more random choices at the beginning of the learning process and next be more "frozen" to guarantee the high convergence speed of the learning process to (semi-)optimal approximation space. In the case of more compound interactions between ag and ENV, e.g., in reinforcement learning, the quality measure Q should be learned using, e.g., awards or penalties received as the results of such interactions. This means that together with searching for an approximation space for the concept it is necessary to search for an approximation space over which the relevant quality measure can be approximated with high quality.

The scheme Sch describes an adaptive strategy ST modifying the induced approximation space AS^* with respect to the changing information about the concept C. To explain this in more detail, let us first assume that a procedure $new_C(ENV, u)$ is given returning from the environment ENV and current information u about the concept C a new piece of information about this concept (e.g., an extension of a sample u of C). In particular, $Inf^{(0)}(C) = new_C(ENV, \emptyset)$ and $Inf^{(k+1)}(C) = new_C(ENV, Inf^{(k)}(C))$ for $k = 0, \ldots$. In Figure 7 $Inf'(C) = Inf^{(1)}(C)$. Next, assuming that operations $Q^{(0)} = Q$, $DS^{(0)} = DS$, $IN^{(0)} = IN$ are given, we define $Q^{(k+1)}$, $DS^{(k+1)}$, $IN^{(k+1)}$,

$DS^{(k+1)}(Inf^{(k+1)}(C))$, and $AS^{*(k+1)}$ for $k = 0, \ldots,$ by

$$(Q^{(k+1)}, DS^{(k+1)}, IN^{(k+1)}) = \tag{22}$$
$$= STR(Q^{(k)}(AS^{*(k)}, Inf^{(k+1)}(C)), Q^{(k)}, IN^{(k)}, DS^{(k)}, AS^{*(k)}, AS^{(k)}_{Inf^{(k)}(C)})$$
$$AS^{(k+1)}_{Inf^{(k+1)}(C)} = DS^{(k+1)}(Inf^{(k+1)}(C)); \quad AS^{*(k+1)} = IN^{(k+1)}(AS^{(k+1)}_{Inf^{(k+1)}(C)}).$$

One can see that the concept of approximation space considered so far should be substituted by a more complex one represented by the scheme Sch making it possible to generate a sequence of approximation spaces $AS^{*(k)}$ for $k = 1, \ldots$ derived in an adaptive process of approximation of the concept C. One can also treat the scheme Sch as a complex information granule [89].

One can easily derive more complex adaptive schemes with metastrategies that make it possible to modify also strategies. In Figure 8 there is presented an

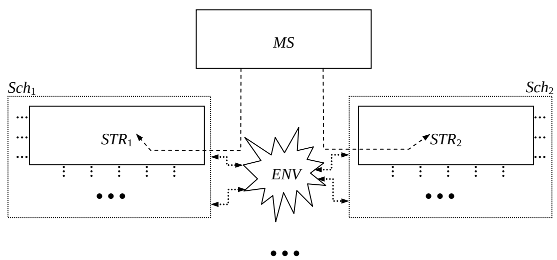


Fig. 8. An example of metastrategy in adaptive concept approximation

idea of a scheme where a metastrategy MS can change adaptively also strategies STR_i in schemes Sch_i for $i = 1, \ldots, n$ where n is the number of schemes. The metastrategy MS can be, e.g., a fusion strategy for classifiers corresponding to different regions of the concept C.

5.3 Adaptation in incremental learning

In incremental learning [21, 31, 35, 64, 73, 104, 108], the main task is to learn the approximation of the function $Q(s, a)$ where s, a denotes a global state of the system and an action performed by an agent ag, respectively, and the real value of $Q(s, a)$ describes the reward for executing the action a in the state s. In approximation of the function $Q(s, a)$, probabilistic models are used. However, for compound real-life problems it may be hard to build such models for such a compound concept as $Q(s, a)$ [113]. In this section, we would like to suggest another approach to approximation of $Q(s, a)$ based on ontology approximation. The approach is based on the assumption that in a dialog with experts an additional knowledge can be acquired making it possible to create a ranking of values $Q(s, a)$ for different actions a in a given state s. We expect that in the explanation given by expert about possible values of $Q(s, a)$ are used concepts from a special ontology of concepts. Next, using this ontology one can follow hierarchical learning methods (see Section 4 and [6–9, 50–52, 85, 86, 89, 90, 105])) to learn approximations of concepts from ontology. Such concepts can have temporal character too. This means that the ranking of actions may depend not only on the actual action and the state but also on actions performed in the past and changes caused by these actions.

Observe that the approximation of domain ontology can be also used in planning (Section 5.4).

5.4 Adaptation and planning

A more compound scheme than what was considered in the previous section can be obtained by considering strategies based on cooperation among the schemes for obtaining concept approximations of high quality. In Figure 9 an adaptive

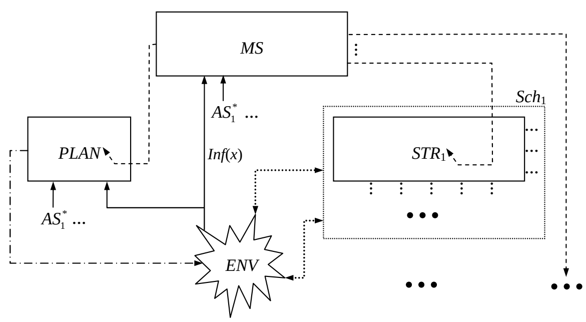


Fig. 9. An example of adaptive plan scheme

scheme for plan modification is presented. $PLAN$ is modified by a metastrategy MS that adaptively changes strategies in schemes Sch_i where $i = 1, \ldots, n$. This is performed on the basis of the derived approximation spaces AS_i^* induced for concepts that are guards (preconditions) of actions in plans and on the basis of information $Inf(x)$ about the state x of the environment ENV. The generated approximation spaces together with the plan structure are adaptively adjusted to make it possible to achieve plan goals.

The discussed example is showing that the context in which sequences of approximation spaces are generated can have complex structure represented by relevant adaptive schemes. The main goal of the agent ag in adaptive planning is to search for approximation of the optimal trajectory of states making it possible for the agent ag to achieve the goal, e.g., to keep as invariants some dependencies between vague concepts. Observe, that searching in adaptive learning for such a trajectory approximation should be performed together with adaptive learning of many other vague concepts which should be approximated, e.g., preconditions for actions, meta actions or plans.

One of the very important issue in adaptive learning is approximation of compound concepts used in reasoning about changes observed in the environment. The agent ag interacting with the environment ENV is recording changes in the satisfiability of concepts from her/his ontology. These changes should be expressed by relevant concepts (features) which are next used for construction of preconditions of actions (or plans) performed by the agent ag. In real-life problems these preconditions are compound concepts. Hence, to approximate such concepts we suggest to use an additional ontology of changes which can be acquired in a dialog with experts. All concepts from the ontology create a hierarchical structure. In this ontology relevant concepts characterizing changes in the satisfiability of concepts from the original ontology are included together with other simpler concepts from which they can be derived. We assume that such an ontology can be acquired in a dialog with experts. Concepts from this ontology are included in the expert explanations consisting of justifications why in some exemplary situations it is necessary to perform some particular actions in a particular order. Next, by approximation of the new ontology (see Section 4 and [6–9, 50–52, 85, 86, 89, 90, 105]) we obtain the approximation of the mentioned above compound concepts relevant for describing changes. This methodology can be used not only for predicting the relevant actions, meta actions or plans but also for the plan reconstruction. In our current projects we are developing the methodology for adaptive planning based on ontology approximation.

6 Conclusions

In the paper we have discussed some problems of adaptive approximation of concepts by agents interacting with environments. These are the fundamental problems in synthesis of intelligent systems. Along this line important research directions perspective arise.

In particular, this paper realizes a step toward developing methods for adaptive maintenance of constraints specified by vague dependencies. Notice that there is a very important problem related to such a maintenance which should be investigated further, i.e., approximation of vague dependencies. The approach to this problem based on construction of arguments "for" and "against" for concepts from conclusions of dependencies on the basis of such arguments from premisses of dependencies will be presented in one of our next paper.

Among interesting topics for further research are also strategies for modeling of networks supporting approximate reasoning in adaptive learning. For example, AR schemes and AR networks (see, e.g., [89]) can be considered as a step toward developing such strategies. Strategies for adaptive revision of such networks and foundations for autonomous systems based on vague concepts are other examples of important issues.

In this paper also some consequences on understanding of vague concepts caused by inductive extensions of approximation spaces and adaptive concept learning are presented. They are showing that in the learning process each temporary approximations, in particular boundary regions are crisp but they are only temporary approximations of the set of borderline cases of the vague concept. Hence, the approach we propose is consistent with the higher order vagueness principle [36].

There are some important consequences of our considerations for research on approximate reasoning about vague concepts. It is not possible to base such reasoning only on *static* models of vague concepts (i.e., approximations of given concepts [59] or membership functions [117] induced from a sample available at a given moment) and on multi-valued logics widely used for reasoning about rough sets or fuzzy sets (see, e.g., [58, 70, 117, 120]). Instead of this there is a need for developing evolving systems of logics which in open and changing environments will make it possible to gradually acquire knowledge about approximated concepts and reason about them. This is related to the view presented by Professor Leslie Valiant (http://people.seas.harvard.edu/ valiant/researchinterests.htm)

> *A specific challenge is to build on the success of machine learning so as to cover broader issues in intelligence. This requires, in particular a reconciliation between two contradictory characteristics – the apparent logical nature of reasoning and the statistical nature of learning.*

Acknowledgment

The research has been supported by the grant NN516 077837, from the Ministry of Science and Higher Education of the Republic of Poland, the National Centre for Research and Development (NCBiR) under grant SP/I/1/77065/10 by the Strategic scientific research and experimental development program: "Interdisciplinary System for Interactive Scientific and Scientific-Technical Information".

References

1. P. Apostoli, A. Kanda (1979). Parts of the continuum: Towards a modern ontology of sciences, Technical Reports in Philosophical Logic, vol. 96 (1). The University of Toronto, Department of Philosophy, Toronto, Canada, 1999. Revised March, 1999.
2. R. M. Axaelrod (1997). *The Complexity of Cooperation*. Princeton University Press, Princeton, NJ.
3. M. Banerjee, M. K. Chakraborty (2003). Foundations of vagueness: a category-theoretic approach. Electronic Notes in Theoretical Computer Science **82**(4): 10-19.
4. J. Bazan (2008). Hierarchical classifiers for complex spatio-temporal concepts. Transactions on Rough Sets IX, Journal Subline, Lecture Notes in Computer Science **5390**: 474–750.
5. Bazan J., Skowron A., Swiniarski R. (2006). Rough sets and vague concept approximation: From sample approximation to adaptive learning. Transactions on Rough Sets V: Journal Subline, Lecture Notes in Computer Science **4100**: 39–62.

6. J. G. Bazan, H. S. Nguyen, J. F. Peters, A. Skowron, M. Szczuka (2003). Rough set approach to pattern extraction from classifiers. In: A. Skowron and M. Szczuka (Eds.), *Proceedings of the Workshop on Rough Sets in Knowledge Discovery and Soft Computing at ETAPS 2003* (RSKD 2003), April 12-13, 2003, Electronic Notes in Computer Science, **82**(4), Elsevier, Amsterdam, Netherlands, 20–29, www.elsevier.nl/locate/entcs/volume82.html.
7. J. G. Bazan, J. F. Peters, A. Skowron (2005). Behavioral pattern identification through rough set modelling. In [103], 688–697.
8. J. Bazan, A. Skowron (2005). On-line elimination of non-relevant parts of complex objects in behavioral pattern identification. In [57], 720–725.
9. J. Bazan, A. Skowron (2005). Classifiers based on approximate reasoning schemes. In B. Dunin-Keplicz, A. Jankowski A. Skowron, M. Szczuka (Eds.), *Monitoring, Security, and Rescue Tasks in Multiagent Systems MSRAS*, Advances in Soft Computing. Springer, Heidelberg, 2005, 191–202.
10. D. Bianucci, G. Cattaneo (2009). Information entropy and granulation coEntropy of partitions and coverings: A summary. Transactions on Rough Sets X, Journal Subline, Lecture Notes in Computer Science **5656**: 15-66.
11. M. Black (1937). Vagueness: An exercise in logical analysis. Philosophy of Science **4**(4): 427–455.
12. A. O. V. Le Blanc (2003). Leśniewski's computative protothetic. *Report (Ph.D. Thesis)*, University of Manchester, Manchester, UK, 2003
13. Z. Bonikowski, U. Wybraniec-Skardowska (2008). Vagueness and roughness. Transactions on Rough Sets IX, Journal Subline, Lecture Notes in Computer Science **5390**: 1–13.
14. L. Breiman (2001). Statistical modeling: The two cultures. Statistical Science **16**(3), 199–231.
15. G. Cantor (1874). Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen. Crelle's Journal für Mathematik **77**: 258–263.
16. G. Cantor (1883). *Grundlagen einer allgemeinen Mannigfaltigkeitslehre*. B. G. Teubner, Leipzig, Germany.
17. R. Casti, A. Varzi (Eds.) (1999). *Parts and Places. The Structures of Spatial Representation.* The MIT Press, Cambridge, MA, 1999.
18. M. K. Chakraborty (2011). On fuzzy sets and rough sets from the perspective of indiscernibility. In *Proceedings of the 4th Indian Conference Logic and Its Applications* (ICLA 2011), Delhi, India, January 5-11, 2011. Lecture Notes in Computer Scence **6521**, Springer, Heidelberg, 22–37.
19. A. Desai (2005). Adaptive complex enterprices. Communications ACM **48**(5), 32–35.
20. T. G. Dietterich (1997). Machine Learning Research: Four Current Directions. AI Magazine **18**(4), 97–136.
21. T. G. Dietterich (2000). Hierarchical reinforcement learning with the MAXQ value function decomposition. Artificial Intelligence **13**, 227–303.
22. D. Dubois, L. Godo, H. Prade, F.Esteva (2005). An information-based discussion of vagueness. In H. Cohen, C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science*, Chap. 40, Elsevier, 892–913.
23. H. Filip (2000). Nominal and verbal semantic structure: Analogies and interactions. Language Sciences **23**: 453–501.
24. K. Fine (1975). Vagueness, truth and logic. Synthese **30**: 265–300.
25. P. Forrest (2002). Sets as mereological tropes. Metaphysical **3**: 5–10.
26. G. Frege (1903). *Grundgesetzen der Arithmetik* vol. 2. Verlag von Hermann Pohle, Jena, Germany.

27. J. Friedman, T. Hastie, R. Tibshirani (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, Heidelberg.
28. M. Ghallab, D. Nau, P. Traverso (2004). *Automated Planning: Theory and Practice*. Elsevier, Morgan Kaufmann, CA.
29. M. Gell-Mann (1994). *The Quark and the Jaguar*. Freeman and Co., NY.
30. Ch. Giraud-Carrier (2000). A note on the utility of incremental learning. AI Communications **13**(4), 215–223.
31. A. McGovern (2002). *Autonomous Discovery of Temporal Abstractions from Interaction with an Environment*. PhD thesis, University of Massachusetts Amherst.
32. P. J. Hoen, K. Tuyls, L, Panait, S. Luke, J.A. La Poutré (2005). An Overview of cooperative and competitive multiagent learning. In K. Tuyls, P. J. Hoen, K. Verbeeck, S. Sen (Eds.), *Learning and Adaption in Multi-Agent Systems: First International Workshop* (LAMAS 2005), Utrecht, The Netherlands, July 25, 2005, 1–46.
33. T. Jech (1997). *Set Theory* (2nd ed). Springer-Verlag, New York.
34. H. Kautz, B. Selman, Y. Jiang (1997). A General Stochastic Approach to Solving Problems with Hard and Soft Constraints. In D. Gu, J. Du, P. Pardalos (Eds.), *The Satisfiability Problem: Theory and Applications*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science **35**, American Mathematical Society, 573–586.
35. L. P. Kaelbling, M. L. Littman, A. W. Moore (1996). Reinforcement learning: A survey. Journal of Artificial Intelligence Research **4**, 237–285.
36. R. Keefe (2000). *Theories of Vagueness.* Cambridge Studies in Philosophy, Cambridge, UK.
37. R. Keefe, P. Smith (1999). *Vagueness: A Reader*. MIT Press, Cambridge, MA, 1999.
38. S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi (1983). Optimization by simulated annealing. Science **220**(4598), 671–680.
39. D. Krause (2003). The mathematics of non-individuality. *Report, Universidade de Sao Paulo*, Sao Paulo, Brazil, 2003.
40. K. Krawiec, R. Słowiński, D. Vanderpooten (1998). Lerning decision rules from similarity based rough approximations. In L. Polkowski, A. Skowron (Eds.), *Rough Sets in Knowledge Discovery 2: Applications, Case Studies and Software Systems*, Physica-Verlag, Heidelberg, 37–54.
41. S. Leśniewski (1929). Grungzüge eines neuen Systems der Grundlagen der Mathematik. Fundamenta Mathematicae **14**: 1–81.
42. M. Luck, P. McBurney, O. Shehory, S. Willmott(2005). *Agent Technology: Computing as Interaction. A Roadmap for Agent-Based Computing*. Agentlink III, the European Coordination Action for Agent-Based Computing, University of Southampton, UK.
43. J. Łukasiewicz (1913). Die logischen Grundlagen der Wahrscheinlichkeitsrechnung, 1913, In: L. Borkowski (Ed.), *Jan Łukasiewicz - Selected Works*, North Holland Publishing Company, Amsterdam, London, Polish Scientific Publishers, Warsaw, 1970, 16–63.
44. Malyszko D., Stepaniuk J.: Adaptive Multilevel Rough Entropy Evolutionary Thresholding, Information Sciences 180(7), 2010, 1138–1158.
45. S. Marcus (1998). The paradox of the heap of grains, in respect to roughness, fuzziness and negligibility. In: Polkowski and Skowron [65], 19–23.
46. L.M. Marx (1999). Adaptive learning and iterated weak dominance. Games and Economic Behavior **26**, 253–278.

47. R. S. Michalski, I. Mozetic, J. Hong, N. Lavrac (1986). The multi-purpose incremental learning system AQ15 and its testing application to three medical domains. In *Proceedings of the 5th National Conference on Artificial Intelligence*, Philadelphia, PA, August 11-15, 1986, Morgan Kaufmann, 1986, 1041–1047.
48. P. Milgrom, J. Roberts (1991). Adaptive and sophisticated learning in normal form games. Games Economic Behavior **3**, 82–100.
49. S. H. Nguyen (2000). Regularity analysis and its applications in Data Mining. In L. Polkowski, T. Y. Lin, S. Tsumoto (Eds.), *Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems*, Physica-Verlag, Heidelberg, Germany, 289–378.
50. S. H. Nguyen, J. Bazan, A. Skowron, H. S. Nguyen (2004). Layered learning for concept synthesis. Transactions on Rough Sets I: Journal Subline, Lecture Notes in Computer Science **3100**: 187–208.
51. T. T. Nguyen, A. Skowron (2003). Rough set approach to domain knowledge approximation. In [115], 221–228.
52. T. T. Nguyen (2005). Eliciting domain knowledge in handwritten digit recognition. In [57], pp. 762–767.
53. E. Orłowska (1984). Semantics of vague conepts. In: G. Dorn, P. Weingartner (Eds.), *Foundation of Logic and Linguistics*, Plenum Press, New York, 465–482.
54. E. Orłowska (1987). Reasoning about vague concepts. Bulletin of the Polish Academy of Sciences, Mathematics **35**: 643–652.
55. S. K. Pal, L. Polkowski, A. Skowron (eds.)(2004). *Rough-Neural Computing: Techniques for Computing with Words*, Cognitive Technologies, Springer-Verlag, Heidelberg.
56. S. K. Pal., B. U. Shankar, P. Mitra (2005). Granular computing, rough entropy and object extraction, Pattern Recognition Letters **26**(16): 2509–2517.
57. S. K. Pal, S. Bandoyopadhay, S. Biswas (Eds.). *Proceedings of the First International Conference on Pattern Recognition and Machine Intelligence* (PReMI'05), December 18-22, 2005, Indian Statistical Institute, Kolkata, Lecture Notes in Computer SCience **3776**, Springer-Verlag, Heidelberg, Germany, 2005.
58. J. Pavelka (1979). On Fuzzy Logic I-III. Zeit. Math Logik Grund. Math. **25**, 45–52, 119-134, 447-464.
59. Z. Pawlak (1991). *Rough Sets: Theoretical Aspects of Reasoning about Data*. System Theory, Knowledge Engineering and Problem Solving **9**, Kluwer Academic Publishers, Dordrecht.
60. Z. Pawlak, A. Skowron (1994). Rough Membership Functions. In R. R. Yager, M. Fedrizzi, J. Kacprzyk (eds.), *Advances in the Dempster-Schafer Theory of Evidence*, John Wiley and Sons, New York, 251–271.
61. Z. Pawlak (1995). Vagueness and uncertainty: a rough set perspective. Computational Intelligence: An International Journal **11**: 217–232.
62. Z. Pawlak (1997). Vagueness - a rough set view. In: Mycielski, J., Rozenberg, G., Salomaa, A. (Eds.) *Structures in Logic and Computer Science*, Lecture Notes in Computer Scence **1261**, Springer, Heidelberg 106-117.
63. Z. Pawlak, A. Skowron (2007). Rudiments of rough sets; Rough sets: Some extensions; Rough sets and boolean reasoning. Information Sciences **177**(1): 3–27; 28–40; 41–73.
64. J. F. Peters, Ch. Henry (2006). Reinforcement learning with approximation spaces. Fundamenta Informaticae **71**(2-3), 323–349.
65. L. Polkowski, A. Skowron (Eds.) (1998). *First International Conference on Rough Sets and Soft Computing* RSCTC'1998, Lecture Notes in Artificial Intelligence **1424**, Springer-Verlag, Warsaw, Poland, 1998.

66. L. Polkowski, T. Y. Lin, S. Tsumoto (Eds.) (2000). *Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems*, Studies in Fuzziness and Soft Computing **56**, Springer-Verlag/Physica-Verlag, Heidelberg, Germany, 2000.
67. L. Polkowski, A. Skowron (2000). Rough mereology in information systems. a case study: Qualitative spatial reasoning. In: Polkowski et al. [66], 89–135.
68. J. F. Peters, A. Skowron (Eds.) (2005). Transactions on Rough Sets III: Journal Subline, Lecture Notes in Computer Science **3400**, Springer, Heidelberg.
69. L. Polkowski (2011). *Approximate Reasoning by Parts. An Introduction to Rough Mereology.* Intelligent Systems Reference Library **20**, Springer, Heidelberg 2011.
70. L. Polkowski (2002). *Rough Sets: Mathematical Foundations.* Physica-Verlag, Heidelberg.
71. L. Polkowski, A. Skowron (1996). Rough mereology: A new pParadigm for approximate reasoning. International Journal of Approximate Reasoning **15**, 333–365.
72. L. Polkowski, A. Skowron (2001). Rough mereological calculi of granules: A rough set approach to computation. Computational Intelligence **17**, 472–492.
73. J. Randlov (2001). *Solving Complex Problems with Reinforcement Learning.* Ph.D. Thesis, University of Copenhagen.
74. S. Read (1995). *Thinking about Logic. An Introduction to the Philosophy of Logic.* Oxford University Press, Oxford, New York.
75. J. Rissanen (1978). Modeling by shortes data description. Automatica **14**, 465–471.
76. J. Rissanen (1985). Minimum-description-length principle. In: S. Kotz, N. Johnson (Eds.), *Encyclopedia of Statistical Sciences*, John Wiley & Sons, New York, NY, pp. 523–527.
77. G. Ronzitti (Ed.) (2011). *Vagueness: A Guide. Logic.* Epistemology, and The Unity of Science Series **19**, Springer, Heidelberg, 2011.
78. B. Russell (1940). *An Inquiry into Meaning and Truth.* George Allen & Unwin Ltd. and W. W. Norton, London and New York, 1940.
79. B. Russell (1903). *The Principles of Mathematics.* George Allen & Unwin Ltd., London, Great Britain, 1903. (2nd Edition in 1937).
80. B. Russell (1923). Vagueness. The Australasian Journal of Psychology and Philosophy **1**: 84–92.
81. S. Shapiro (2006). *Vagueness in Context.* Clarendon Press, Oxford, 2006.
82. P. Simons (Ed.) (1987). *A Study in Ontology.* Oxford University Press, Oxford, UK, 1987.
83. A. Skowron (2000). Rough Sets in KDD. *16-th World Computer Congress* (IFIP'2000), Beijing, August 19-25, 2000, In Z. Shi, B. Faltings, M. Musen (Eds.), *Proceedings of Conference on Intelligent Information Processing* (IIP2000), Publishing House of Electronic Industry, Beijing, 2000, 1–17.
84. A. Skowron (2001). Toward intelligent systems: Calculi of information granules, (keynote talk at RSTGC'01), Matsue, May 20-22, Japan, In: S. Hirano, M. Inuiguchi, and S. Tsumoto (Eds.), *Proc. Int. Workshop on Rough Set Theory and Granular Computing* (RSTGC'2001), May 20-22, 2001, Matsue, Japan, Bull. Int. Rough Set Society 5(1-2), 2001, 9-30; see also: Terano T., Nishida T., Namatame A., Tsumoto S., Ohsawa Y., Washio T.(Eds.), *New Frontiers in Artificial Intelligence, Joint JSAI Workshop Post Proceedings*, Lecture Notes in Artificial Intelligence **2253**, Springer, Heidelberg, 2001, 251–260.
85. A. Skowron (2004). Approximate reasoning in distributed environments. In N. Zhong, J. Liu (Eds.), *Intelligent Technologies for Information Analysis*, Springer, Heidelberg, 433–474.

86. A. Skowron, J. Peters (2003). Rough sets: Trends and challenges. In [115], 25–34.
87. A. Skowron. Rough sets and vague concepts. Fundamenta Informaticae 64(1-4) (2005) 417–431.
88. A. Skowron, J. Stepaniuk (1996). Tolerance Approximation Spaces. *Fundamenta Informaticae* **27**, 245–253.
89. A. Skowron, J. Stepaniuk (2004). Information granules and rough-neural computing. In [55], 43–84.
90. A. Skowron, J. Stepaniuk (2005). Ontological framework for approximation. In [102], 718–727.
91. Skowron A., Stepaniuk J.: Tolerance approximation spaces. Fundamenta Informaticae 27, 1996, 245–253.
92. A. Skowron, J. Stepaniuk, R. Swiniarski: Rough granular computing based on approximation spaces. Information Sciences, 2011 (in print).
93. A. Skowron, R. Swiniarski (2005). Rough sets and higher order vagueness. In [102], 33–42.
94. A. Skowron, R. Swiniarski, P. Synak (2004). Approximation spaces and information granulation. In [68], 175–189.
95. A. Skowron, M. Szczuka (2010). Toward interactive computations: A rough-granular approach. In J. Koronacki, S. Wierzchon, Z. Ras, J. Kacprzyk (Eds.), *Advances in Machine Learning II, Dedicated to the Memory of Professor Ryszard Michalski, Studies in Computational Intelligence* **263**, Springer, Heidelberg, 23-42.
96. A. Skowron, P. Wasilewski (2011). Information systems in modeling interactive computations on granules, Theoretical Computer Science, (2011), doi:10.1016/j.tcs.2011.05.045
97. B. Smith (1995). Formal ontology, common sense and cognitive science. International Journal of Human-Computer Studies **43**: 641–667.
98. J. G. Stell (2000). Boolean connection algebras: A new approach to the region-connection calculus. Artificial Intelligence **122**: 111–136.
99. J. Stepaniuk, J. Bazan, A. Skowron (2005). Modelling complex patterns by information systems. Fundamenta Informaticae **67**(1-3), 203–217.
100. D. Ślęzak (2002). Approximate entropy reducts. Fundamenta Informaticae **53**, 365–387.
101. D. Ślęzak, P. Wasilewski (2007). Foundations of rough sets from vagueness perspective. In A. E. Hassanien, Z. Suraj, D. Ślęzak, P. Lingras (Eds.), *Rough Computing: Theories, Technologies and Applications*, IGI Global, Hershey , PA, 1–37.
102. D. Ślęzak, G. Wang, M. Szczuka, I. Düntsch, Y. Yao (Eds.) (2005). *Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing* (RSFDGrC'2005), Regina, Canada, August 31-September 3, 2005, Part I, Lecture Notes in Artificial Intelligence, vol. 3641. Springer-Verlag, Heidelberg, Germany.
103. D. Ślęzak, J. T. Yao, J. F. Peters, W. Ziarko, X. Hu (Eds.) (2005). *Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing* (RSFDGrC'2005), Regina, Canada, August 31-September 3, 2005, Part II, Lecture Notes in Artificial Intelligence, vol. 3642. Springer-Verlag, Heidelberg, Germany.
104. J. Stepaniuk, A. Skowron, J. Peters, R. Swiniarski (2006). Calculi of approximation spaces. Fundamenta Informaticae **72**(1-3) 363–378.
105. P. Stone (2000). *Layered Learning in Multi-Agent Systems: A Winning Approach to Robotic Soccer.* The MIT Press, Cambridge, MA.
106. R. Sun (ed.) (2006). *Cognition and Multi-Agent Interaction. From Cognitive Modeling to Social Simulation.* Cambridge University Press, New York, NY.

107. Z. Suraj (1998). The synthesis problem of concurrent systems specified by dynamic information systems. In L. Polkowski, A. Skowron (Eds.), *Rough Sets in Knowledge Discovery 2: Applications, Case Studies and Software Systems*, Physica-Verlag, Heidelberg, 418–448.
108. R. S. Sutton, A. G. Barto (1998). *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA.
109. P. Utgoff (1988). An incremental ID3. In *Proceedings of the Fifth International Workshop on Machine Learning*, Ann Arbor, Michigan, June 12-14, 1988, Morgan Kaufmann, 107–120.
110. D. Vakarelov, (1987). Abstract characterization of some knowledge representation systems and the logic NIL of nondeterministic information. In Ph. Jorrand, V. Sgurev (Eds.), *Artificial Intelligence II*, North-Holland, 255-260.
111. W. Van Wezel, R. Jorna, A. Meystel (2006). *Planning in Intelligent Systems: Aspects, Motivations, and Methods*. John Wiley & Sons, Hoboken, New Jersey.
112. A. C. Varzi (2005). Change, temporal parts, and the argument from vagueness. Dialectica **59**(4): 485–498.
113. V. Vapnik (1998). *Statistical Learning Theory.* John Wiley & Sons, New York, NY.
114. P. Vopenka (1979). *Mathematics in the Alternative Set Theory*. Teubner, Leipzig, 1979.
115. G. Wang, Q. Liu, Y. Yao, A. Skowron (Eds.) (2003). *Proceedings of the 9-th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing* (RSFDGrC'2003), Chongqing, China, May 26-29, 2003, Lecture Notes in Artificial Intelligence **2639**, Springer-Verlag, Heidelberg, Germany, 2003.
116. A. Wojna (2001). Constraint based incremental learning of classification rules. In [122], 428-435.
117. L. A. Zadeh (1965). Fuzzy sets. Information and Control **8**, 333–353.
118. L. A. Zadeh (1973). Outline of a new approach to the analysis of complex systems and decision processes. Trans. on Systems, Man and Cybernetics **SMC-3**, 28-44.
119. L. A. Zadeh (1975). The concept of a linguistic variable and its application to approximate reasoning I, II, III, Information Sciences **8**, 199-257, 301-357; **9** 43-80.
120. L. A. Zadeh (1996). Fuzzy Logic = Computing with Words. IEEE Transactions on Fuzzy Systems **2**, 103–111.
121. W. Ziarko (1993). Variable precision rough set model. Journal of Computer and System Sciences **46**, 39–59.
122. W. Ziarko, Y. Yao (Eds.) (2001). *Proceedings of the 2nd International Conference on Rough Sets and Current Trends in Computing (RSCTC'2000)*, Banff, Canada, October 16-19, 2000, Lecture Notes in Artificial Intelligence **2005**, Springer-Verlag, Heidelberg, Germany, 2001.

Jan Bazan
Chair of Computer Science, University of Rzeszów,
Rejtana 16A, 35-959 Rzeszów, Poland.
`bazan@univ.rzeszow.pl`

Andrzej Skowron
Institute of Mathematics, Warsaw University,
Banacha 2, 02-097 Warsaw, Poland.
`skowron@mimuw.edu.pl`

Jarosław Stepaniuk
Department of Computer Science, Białystok University of Technology,
Wiejska 45A, 15-351 Białystok, Poland.
j.stepaniuk@pb.edu.pl

Roman Swiniarski
Institute of Computer Science, Polish Academy of Sciences,
Ordona 21, 01-237 Warsaw, Poland.
and
Department of Mathematical and Computer Sciences, San Diego State University,
5500 Campanile Drive San Diego, CA 92182, USA.
rswiniar@sciences.sdsu.edu

Enigma of Contradiction

Lopamudra Choudhury

Abstract. The Law of Non-Contradiction has played a very crucial role in the history of reasoning. Irrespective of Indian, Chinese, Greek or Western culture, beginning from Sanjaya (pre Buddhist philosopher) to the contemporary logicians the law of non-contradiction has preoccupied thinkers over the ages. Philosophers, mathematicians, information scientists, logicians, and poets alike have been baffled by the enigmatic principle and tried to reconcile it within their own enterprise. This paper attempts to provide a broad overview of some interpretations of principle and tries to link its significance with the stand point of cognitive science.

1.1 Aristotle's view of non-contradiction

Aristotle considers the principle of non-contradiction as the firmest of his first principle or basics of his theoretical standpoint which holds an important place in his first philosophy i.e metaphysics. This principle is prior to any theoretical construct irrespective of the subject matter of any science. Aristotle considers the principle of non-contradiction to be the principle of scientific enquiry, reasoning and communication. Communication and for that matter in any informal content the law of non-contradiction is presupposed. It a presupposition of human reasoning. Unlike modus ponens it is not a rule of inference. To Aristotle it is the principle which "is necessary for anyone to have who knows any of the things that are"[1]. It is also not a hypothesis.

Aristotle says that it cannot be demonstrated. Although it is not susceptible to demonstration, it is subject to "*elenctic refutation*" where by he means the Socratic method of argument. By the use of "*elenchus*" Socrates gets his opponent to refute himself out of his own mouth. In such a situation the opponent gives up one of his claims to be consistent. This is as a matter of fact the *reduction ad absurdum* method of argumentation.

Aristotle discusses three versions of the principle of non-contradiction.

The first is an ontological version, which reads as "It is impossible for the same thing to belong and not to belong at the same time to the same thing and in the same respect."[2]

---

[1] Metephysics IV3 1005b15

[2] Metaphysics IV 3 1005b19-20.

The second version, a doxastic one is as follows: "It is impossible to hold (suppose) the same thing to be and not to be"[3]. This can be understood as the claim that it is impossible to hold the same thing to be F and not to be F. Taking recourse to psychology where one can claim that it is possible to hold contradictory belief, the formulation may be revised as, one cannot believe that x is F and x is not F.

The third is semantic version which says "opposite assertions cannot be true at the same time"[4].this version is neutral about the internal structure of the assertion. However, Aristotle assumes that any assertion involves predicating one thing of the other.

If we accept Aristotelian essentialism then we have to accept the view that individuals belonging to natural kinds have essential natures that are definable. There cannot be any change in the essence of the individual but change in their accidental property is accepted. A human being in essence is a human being and he cannot undergo change to become an animal, on the other hand, a human being can undergo change by changing his fashion. This may tell on the identity of the human being but not on its specieshood. One cannot be both human and non- human but there is no absurdity or contradiction in being both traditional and modern at the same time. At this juncture, it would be interesting to bring in the views of Anaxagorous and Democritus who hold that contradictions are true at the same time; they see contraries come into existence out of the same thing. Aristotle distinguishes between the actual and the potential to resolve this issue. An object can be potentially P and potentially not P, but it cannot actually be P and not P at the same time. In response to Heraclitean doctrine of flux, that one cannot step into the same river twice, Aristotle maintains that there is no radical flux. Though the water is flowing constantly, the river is still identifiable.

1.2 Contradiction in logic

In Western logic, encountering a contradiction in a proof empowers one to jump into any conclusion in a two valued logic. This is why the law is popularly known as the Law of Explosive Premises. However, the explosion is essentially aided by the Law of Addition (unrestricted). It is unrestricted liberality of the law of addition that can be traced as the cause of the trouble. It is raining and it is sunny are not contradictory. They are the phenomena that are seldom found to go together. It is raining and it is not raining are contradictory to each other. Contradictory situations are not instantiated. No two facts are contradictory but the corresponding statements are.

The addition of any q given p needs to be constrained. If reasoning is rooted in commonsensical argumentation, then the inference from the law of absurdity seems to be counter - intuitive. Whereas the law of addition is not counter-intuitive.

---

[33] Ibid, IV 3 1005b24 cf. 1005b 29-30
[4] Ibid, IV6 1011b13-20.

1.3 Paraconsistent logics

It is important to mention that negation is a syntactic concept whereas contradiction is a semantic concept. It is the interpretation of negation in various contexts that either give rise to contradiction (as in the classical case) or does not give rise to contradiction (as in case of many valued logics). Change in the concept of negation may help in doing away with contradiction. Syntactically it would mean these are different negations expressed through different symbols. Many valued and paraconsistent logics are some examples to cite in this regard. It is the demand of consistent theories that they cannot hold contradictory propositions as thesis. Many valued logics and paraconsistent logics take different paths to solve the problem. The paraconsistent logicians primarily take into consideration the concept of contradiction understood as pairs of propositions which cannot be true together. To them one contradiction need not lead to an explosion where anything goes. In many valued systems there are no contradictions as there are values more than two. There are values beyond truth and falsity.

1.4 Open universe

In place of the classical set theory where the universe is closed by definition, if we bring in the notion of an open universe[5], we cannot say that either x belongs to P or x does not belong to P. For example, in classical set theoretic assertion, we say for any individual x, x belongs to some predicate P or x does belong to the complement of P. However, in a system with open universe, from x does not belong to P we cannot say x belongs to the complement of P but we say, know not where x is. The system is so formulated that expressions of the form x is in P and x is not in P will not be accommodated.

2.1 Jaina logical system

In the Jaina logical system p.~ p is interpreted like that of the many valued systems. Here, discontent with the two- valued situation, a modified notion of negation is brought in. The emphasis now shifts to the meaning of negation. The Jainas are of the opinion that there can be reasons for affirming or denying propositions; thus its theory of relativity of truth was formulated. The doctrine of standpoints (*nayavāda*) and the doctrine of conditioned predication (*syādavāda*) were formulated to justify every statement is true from one standpoint and false from another; thus both S is P and S is not P can be true from different standpoints. This formulation bypasses the law of noncontradiction in its own way. Because of this no statement can be absolutely true, it has only relative truth.

The sevenfold conditional predication of the Jainas are the following:

1. It is possible "S is P"
2. It is possible "S is not-P"

---

[5] Choudhury L and Chakraborty M, *On Representing Open Universe*, Studies in logic, China, forthcoming issue, 2012.

3. It is possible "S is both P and not-P"
4. It is possible "S is indescribable"
5. It is possible "S is both P and indescribable"
6. It is possible "S is not-P and indescribable"
7. It is possible "S is both P and not-P and also indescribable"

The meaning of indescribability is that S is neither P nor not-P. P and not P are opposites and therefore cannot be applied to the same subject. In this situation, we can say about S that no predicate can be referred to it; i.e. S is neither P nor not-P. If there is a contradiction at all, then there will not be any predication.

2.2 Buddhist logic

It will be relevant to bring in the four-fold negation of the Indian logicians and metaphysicians. The principle of four cornered negation of the Indian logicians, P.T Raju[6] holds, is proposed by Sanjaya, a predecessor of Buddha, may be of relevance here:

"S is neither P nor not-P, nor both P and not-P, nor neither P nor not-P"; clearly in this metaphysics one withholds ones position. It was not possible to fix position either for approval or for rejection of a particular standpoint. The proponents of four cornered negation would criticize any view positive or negative but would themselves not hold any. Even Buddha at times refused to answer questions about what he called the imponderables, such as "Is there an eternal entity like the self?" Though some of his followers interpreted his silence as denial and developed their own theory, some considered his silence to be refusal to say "yes" or "no". The reason for Buddha's silence lay in the indefiniteness of the nature of the concept of the *self* or *atman*. Nagarjuna, one of the renowned Buddhist philosophers, adopted the same principle to prove his doctrine of *sunya* or void. To him the reality is neither being, nor non-being, nor both nor neither. It is interesting to note that what was for Sanjaya a principle or method of doubt or skepticism, became a metaphysical principle in the hands of Nagarjuna to express the ultimate reality. It is interesting to note that the supporters of the four cornered negation transcend the principle of non contradiction. P and not-P situation do not arise here.

2.3 Navya-nyāya view

According to Sibajiban Bhattacharyya[7] as far as the relation between cognition of contradictory proposition is concerned, the problem reduces to the fact as to whether it is logically possible for one and the same person to believe that p and believe that not-p. Navya-nyāya solves this problem by distinguishing between actual state of believing and the disposition to believe. We may recall the Aristotelian distinction between the actual and the potential here. In case of the cognition of two contradictory propositions the presence of one act prevents the origination of the other act, which

[6] P.T.Raju, The Principle of four cornered negation in Indian Philosophy, The review of Metaphysica, Vol 7, No.4, (1954), pp 694-713.

[7] Sibajiban Bhattacharyya, Some features of Navya nyaya logic, Philosophy East and West, Vol 24, No 3 (1974) pp 329-342.

stops them from being present at the same moment. There is a fundamental distinction between disposition and an act. So long as part of conscious act one believes that p, it will prevent a person from believing that not-p; however as dispositions the belief that p and the belief that not-p both can be believed provided the beliefs are not recalled. "A person cannot *consciously*, *knowingly*, hold self-contradictory beliefs"[8]. As unconscious dispositions one can hold contradictory beliefs.

The law of non-contradiction can be stated in the following way: S cannot be both P and not-P. S is both P and not-P is self-contradictory. The Navya-nyāya replaces the 'is' by the expression 'related to'. The law of non-contradiction can then be reformulated as 'S cannot be related to t and absence of t by the same occurrence exacting relation'.

Bhattacharyya further explains by arguing that we can say that the pen is on the table which means it is in contact with the surface of the table; but it will not make any sense to say that the negation of the above proposition, absence - of - the – pen, is in contact with the surface of the table; if we say that absence of pen on the table, it would mean absence of pen is in contact with the surface of the table, which is meaningless for this violates the definition of 'contact'. If the pen cannot both be and not be on the table, then we shall have to say that 'the surface of the table is in contact with the pen' will be self contradictory with 'the surface of the table is in some other specific relation with the absence -of - the - pen'. There is no other way of specifying the nature of the 'some other relation'.

3.1 Chinese philosophy

Dialectical thinking is a form of folk wisdom in Chinese culture. The Chinese have a preference for proverbs containing seeming contradiction. It is interesting to note that in the version of Peng and Nisbett when two contradictory propositions are presented, the Americans polarized their views and the Chinese moderately accept both the propositions.

Depending on the nature of reality the metaphysical conjectures are formed. The Chinese philosophy hold reality as a process-- life is constantly changing from one state of being to another. It reminds us of the Heraclitean view that one cannot step into the same river twice.

The principle of contradiction (*Mao Dun Lu*) states that the reality is full of contradictions. Just as there is constant change, there is constant contradiction: old and new, strong and weak, day and night, etc. The world is a single entity integrated over opposites. According to Lao-zi, the founder of Daoist school, "when the people of the world all know beauty as beauty, there arises the recognition of ugliness; when all know the good as good, there arises recognition of evil. And so being and nonbeing produce each other.."[9] . The Daoist hold that the two sides of any contradiction exist in harmony, opposed but connected and more importantly mutually connecting.

---

[8]Ibid, p.332

[9] Lao-Zi (570-490BC/1993) *The Book of Lao Zi* Beijing, Foreign Language Press. P.16

3.2 Chakraborty's incorporation of freedom

Nothing is isolated, neither are the polar concepts. The concepts belong to the extreme ends of an imaginary line which are placed at the ends of the spectrum. I would like to have both the spectrum in my world; I would like to have light and darkness, good and bad, flower and weed in my world. Without the polar concepts the reality would be incomplete. Chakraborty[10] elegantly brings in the notion of freedom to construct ones own world by whatever concepts one wishes. There is individual choice to built and rebuilt one's conceptual world just as the painter or the poet does engage in creation by deconstructing and reconstructing forms or thought at his own whim. The contradiction in the conceptual world is different from the contradiction of propositions.

3.3 Contradiction and cognitive science

Categorization is an essential prerequisite to perception and knowledge. Our knowledge formation would be impossible if we do not categorize to identify entities. Our categorization as part of cognition and Aristotle's talk about essence are somewhat related. A child looks confused if he is asked to bring a flower and not to bring a flower at the same time or taught someone to be a teacher and not a teacher at the same time. No knowledge of the principle is needed to restrain the child from the execution of an act. That some animal is a cow is taught by contrasting it with non-cows. Inconsistency in the request or teaching, in the above example, withdraws the child from engaging in any action. Towards the extreme ends of the spectrum, the polar concepts belong to two different categories which are contradictory to each other; we cite different examples to understand different polar concepts and at times we take the help of contrasting concepts to understand one concept. A child is taught the concept of good by contrasting it with the concept of bad by taking appropriate examples. Thus at the cognitive level contradiction is welcome as an aid to concept forming operations.

3.4 Conclusion

It is usually accepted that contradiction leads to inconsistency. However, even consistency is questioned by Chakraborty in some practical or theoretical enterprise. Here too individual choice plays important role in preserving or not preserving consistency of the domain of discourse. However, consistency and its elegant properties cannot be outright rejected. Some situations may embrace paraconsistency but paraconsistency cannot penetrate into every mode of reasoning. We have seen through the itinerary of our exploration of contradiction, how varied and versatile role it plays in the human conceptual framework.

[10] Mihir Chakraborty(2011), *Achin pakhir ghar* in Ganiter Dharapat o Gappasappa, , Nandimukh Sansad, pp152-156.

References

1. Barnes, J. The Complete Works of Aristotle: The Revised Oxford Translation. Princeton: Princeton University Press, vols 1 and 2. (1984)
2. Bhattacharyya Sibajiban, Some features of Navya nyāya logic, Philosophy East and West, Vol24, No 3 (1974) 329-342
3. Choudhury L and Chakraborty M, On Representing Open Universe, Studies in Logic, China, forthcoming (2012)
4. Dutta Soma and Chakraborty, M.K Negation and Paraconsistent Logics, Log. Univers, Vol 5, pp167-176, Springer Basel AG (2011)
5. Gottlieb, Paula: Aristotle on Non-contradiction, in The Stanford Encyclopedia of Philosophy (Summer 2011 Edition), Edward N. Zalta (ed.), URL =<http://plato.stanford.edu/archives/sum2011/entries/ aristotle-noncontradiction/>.(2011)
6. Jaeger, W.: Aristotelis Metaphysica, Oxford Classical Text. Oxford: Clarendon Press. (1951)
7. Lao-Zi: The Book of Lao Zi Beijing, Foreign Language Press. (570?-490?BC/1993)
8. Mihir Chakraborty: Achin pakhir ghar (in Bangla) in Ganiter Dharapat o Gappasappa, Nandimukh Sansad, (2011)152-156
9. Nicholas Rescher: Topics in Philosophical logic, D. Reidal Pub Company, Dordrechet- Holland(1968)
10. Peng Kaiping, Nisbett R.E: Culture, Dialectics and Reasoning about Contradiction, Dissertation submitted by Peng under the guidance of Richard Nisbett, University Michigan
11. Raju P.T: The Principle of fur cornered negation in Indian Philosophy, The Review of Metaphysica, Vol 7, No.4, (1954) 694-713

Lopamudra Choudhury
Department of Philosophy and School of Cognitive Science,
Jadavpur University, Kolkata, India.
choudhuryl@yahoo.com

Some technical features of the graded consequence

Cristina Coppola and Giangiacomo Gerla

1 Introduction

This paper is devoted to examine some mathematical features of Chakraborty's theory of graded consequence (see [3],[4],[5],[6]). Namely we emphasize the suitability of analyzing the connections of such a fundamental approach to fuzzy logic with the notions of canonical extension of a deduction apparatus, closure operator, compactness, recursive enumerability (see [1],[2],[8],[9],[10]).

2 Preliminaries on fuzzy logic

We denote by U the interval $[0,1]$ and we look this interval as a complete lattice in which $\lambda \wedge \mu = \inf\{\lambda, \mu\}$ and $\lambda \vee \mu = \sup\{\lambda, \mu\}$. Given a nonempty set S we call *fuzzy subset* of S any map $s : S \to U$. The class U^S of all fuzzy subsets of S defines a complete lattice whose join and meet operations we call *union* and *intersection*, respectively. We define the *complement* $-s$ of s by setting $-s(x) = 1 - s(x)$ for every $x \in S$. Let's call *continuous chain* an order-reversing family $(S_\lambda)_{\lambda \in U}$ of subsets of S such that $S_\mu = \cap_{\lambda<\mu} S_\lambda$. Then we can identify the fuzzy subsets of S with the continuous chains of subsets of S. Indeed, every fuzzy subset s is associated with the continuous chain $C(s,\lambda))_{\lambda \in U}$ of its cuts, where $C(s,\lambda) = \{x \in S : s(x) \geq \lambda\}$. Since for every $x \in S$

$$s(x) = \sup\{\lambda \in U : x \in C(s,\lambda)\},$$

such a correspondence is injective. Conversely, given any continuous chain $(S_\lambda)_{\lambda \in U}$ of subsets of S, define s by setting

$$s(x) = \sup\{\lambda \in U : x \in S_\lambda\}.$$

Then s is a fuzzy subset whose family of cuts coincides with $(S_\lambda)_{\lambda \in U}$. This proves that the correspondence is one-to-one.

Let F be a set whose elements we call *formulas*, then an *Hilbert deduction system*, in brief an H-system, is a pair $\Sigma = (LA, IR)$ such that LA is a subset of F, *the set of logical axioms*, and IR a set of inference rules. In turn, an *inference rule* is a partially defined n-ary map $r : F^n \to F$. We denote by $Dom(r)$ the domain of r. Given $X \subseteq F$, a *proof* π *of a formula* α *under the hypotheses* X is any sequence $\alpha_1, ..., \alpha_m$ of formulas such that $\alpha_m = \alpha$ and, for any $i = 1, ..., m$,

- either $\alpha_i \in X$ (α is a hypothesis)
- or $\alpha_i \in LA$ (α is an axiom)

- or $\alpha_i = r(\alpha_{s(1)}, ..., \alpha_{s(n)})$, where $r \in IR$ and $s(1) < i, ..., s(n) < i$ (α is obtained by an inference rule).

Given a set X of formulas, we write $X \vdash \alpha$ to denote that a proof of α exists whose hypotheses are contained in X. The *deduction operator* $D : P(F) \to P(F)$ *associated with* Σ is defined by setting, for any set X of formulas,

$$D(X) = \{\alpha \in F : X \vdash \alpha\}. \tag{1}$$

A *theory* is a set T of formulas such that $D(T) = T$. This is equivalent to say that T contains the set LA of logical axioms and that T is closed with respect to the inference rules, i.e., for every n-ary inference rule r,

$$\alpha_1, ..., \alpha_n \in T \text{ and } (\alpha_1, ..., \alpha_n) \in Dom(r) \Rightarrow r(\alpha_1, ..., \alpha_n) \in T.$$

By extending the above definitions, we define a *fuzzy Hilbert system*, in brief *fuzzy H-system*, a pair $\Sigma = (la, IR)$ where la is a fuzzy subset of F, the *fuzzy subset of logical axioms*, and IR is a set of fuzzy inference rules. In turn, a *fuzzy inference rule* is a pair $r = (r', r'')$, where

- r' is a partial n-ary operation on F whose domain we denote by $Dom(r)$;
- r'' is a join-preserving n-ary operation on U, i.e.

In other words, an inference rule r consists

- of a syntactical component r' that operates on formulas (in fact, it is an inference rule in the usual sense);

- of a valuation component r'' which operates on truth values to calculate how the truth value of the conclusion depends on the truth values of the premises.

We indicate an application of an inference rule r by

$$\frac{\alpha_1, ...\alpha_n}{r'(\alpha_1, ...\alpha_n)} \quad ; \quad \frac{\lambda_1, ..., \lambda_n}{r''(\lambda_1, ..., \lambda_n)}.$$

A *proof* π of a formula α is a sequence $\alpha_1, ..., \alpha_m$ of formulas such that $\alpha_m = \alpha$, together with a sequence of related '*justifications*'. This means that, given any formula α_i, we must specify whether

(i) α_i is assumed as a logical axiom; or
(ii) α_i is assumed as a proper axiom; or
(iii) α_i is obtained by a rule (in this case we must indicate also the rule and the formulas from $\alpha_1, ..., \alpha_{i-1}$ used to obtain α_i).

The *valuation* $Val(\pi, v)$ *of* π *with respect to* v is defined by induction on the length m of π as follows. If α_m is assumed as a logical axiom, we set $Val(\pi, v) = la(\alpha_m)$ if α_m is assumed as a hypothesis, we set $Val(\pi, v) = v(\alpha_m)$. Finally, if α_m is obtained by the inference rule $(r^{'}, r^{''})$, i.e. $\alpha_m = r'(\alpha_{s(1)}, ..., \alpha_{s(n)})$, where $s(1) < m, ..., s(n) < m$, then we set

$$Val(\pi, v) = r^{''}(Val(\pi_{s(1)}, v), ..., Val(\pi_{s(n)}, v)).$$

If α is the formula proven by π, the meaning we assign to $Val(\pi, v)$ is that: *given the information v, π assures that α holds at least with degree $Val(\pi, v)$.*

Finally, we define the *deduction operator* $D : [0,1]^F \to [0,1]^F$ by setting

$$D(v)(\alpha) = \sup\{Val(\pi, v) : \pi \text{ is a proof of } \alpha\}. \tag{2}$$

A *fuzzy theory* is a fuzzy subset τ of formulas such that $D(\tau) = \tau$. Equivalently, τ is a fuzzy theory provided that it contains the fuzzy subset la of logical axioms and it is closed with respect to the fuzzy inference rules, i.e. for every n-ary inference rule $r = (r', r'')$,

$$\tau(r'(\alpha_1, ..., \alpha_n)) \geq r''(\tau(\alpha_1), ..., \tau(\alpha_n)).$$

It is immediate to prove that, given any fuzzy subset of hypothesis v, $D(v)$ is a fuzzy theory, the least fuzzy theory containing v.

There is no difficulty to extend these definitions in such a way that infinitary inference rules are admitted.

3 Graded consequence relations as theories of a fuzzy deduction apparatus

We recall the main definitions in Chakraborty's theory. If F is the set of formulas of a logical language, we set $SEQ = P(F) \times F$ and we call *sequents* the elements in SEQ. Then a sequent is a pair (X, α) where X is a set of formulas and α a formula. The intended meaning of (X, α) is the metalogical claim "α is a consequence of X". We call *sets-formulas relation*, in brief *sf-relation*, a relation from $P(F)$ to F, i.e. a subset of SEQ. Also, we denote by $\vdash$ such a kind of relation, we write $X \vdash \alpha$ instead of $(X, \alpha) \in \vdash$ and, given $Z \in P(F)$, we write $X \vdash Z$ to denote that $X \vdash \alpha$ for every $\alpha \in Z$. A *sf*-relation $\vdash$ is a *consequence relation* if

(j) $X \vdash \alpha$ whenever $\alpha \in X$,
(jj) $X \vdash \alpha \Rightarrow X \cup Y \vdash \alpha$,
(jjj) $X \vdash Z$ and $X \cup Z \vdash \alpha \Rightarrow X \vdash \alpha$.

If $\vdash$ is a consequence relation and $X \vdash \alpha$, then α *is a consequence of* X.

By extending such a definition, we interpret any fuzzy subset of sequents, as a fuzzy relation $g : SEQ \to [0,1]$ from $P(F)$ to F and we call it *graded sf-relation*. If X is a set of formulas and α is a formula, we write $g(X \vdash \alpha)$ instead of $g((X, \alpha))$. Also, we set, for every X and Z in $P(F)$,

$$g(X \vdash Z) = \inf\{g(X \vdash z) : z \in Z\}. \tag{3}$$

Definition 3.1. *We say that g is a* graded consequence *if, for every X, Y, $Z \in P(F)$ and $\alpha \in F$,*

(i) *$g(X \vdash \alpha) = 1$ for every $\alpha \in X$,*

(ii) $g(X \cup Y \vdash \alpha) \geq g(X \vdash \alpha)$,
(iii) $g(X \vdash \alpha) \geq g(X \vdash Z) \wedge g(X \cup Z \vdash \alpha)$.

If $\lambda = g(X \vdash \alpha)$ we say that α *is a consequence of* X *with degree* λ. Given a graded *sf*-relation $g : SEQ \to U$, we denote by $\vdash_\lambda$ its λ-cut $C(g, \lambda) = \{(X, \alpha) \in SEQ : g(X, \alpha) \geq \lambda\}$.

Theorem 3.1. *The following are equivalent:*

(a) *$g : SEQ \to U$ is a graded consequence.*
(b) *every cut $\vdash_\lambda$ is a consequence relation.*

Proof. $(a) \Rightarrow (b)$. It is self-evident that $\vdash_\lambda$ satisfies (j) and (jj). To prove (jjj), suppose $X \cup Z \vdash_\lambda \alpha$ and that $X \vdash_\lambda z$ for every $z \in Z$. Then $g(X \cup Z \vdash \alpha) \geq \lambda$ and $g(X \vdash z) \geq \lambda$ for every $z \in Z$. Consequently, $g(X \vdash Z) \geq \lambda$ and, by (iii) of Definition 3.1, $g(X \vdash \alpha) \geq \lambda$. Hence, $X \vdash_\lambda \alpha$.
$(b) \Rightarrow (a)$. Let X be a set of formulas and $\alpha \in X$. Then, since $\vdash_1$ is a consequence relation, $X \vdash_1 \alpha$, i.e., $g(X \vdash \alpha) = 1$. Let Y be a set of formulas and $\lambda = g(X \vdash \alpha)$. Then, since $X \vdash_\lambda \alpha$, we have $X \cup Y \vdash_\lambda \alpha$, and therefore, $g(X \cup Y \vdash \alpha) \geq \lambda = g(X \vdash \alpha)$. Finally, given any set Z of formulas, set $\lambda = g(X \vdash Z) \wedge g(X \cup Z \vdash \alpha)$. Then, since $\vdash_\lambda$ is a consequence relation, $X \vdash_\lambda z$ and for every $z \in Z$, $X \cup Z \vdash_\lambda \alpha$, we may conclude that $X \vdash_\lambda \alpha$. Thus $g(X \vdash \alpha) \geq g(X \vdash Z) \wedge g(X \cup Z \vdash \alpha)$. The remaining part of the theorem is evident. □

Such a theorem enables us to prove the following way to represent the graded consequences.

Corollary 3.1. *A graded sf-relation g is a graded consequence iff a continuous family $(\vdash_\lambda)_{\lambda \in U}$ of consequence relations exists such that*

$$g(X \vdash \alpha) = \sup\{\lambda \in U : X \vdash_\lambda \alpha\}. \tag{4}$$

Now, consider the Hilbert deduction system $\Sigma = (LA, IR)$, which we call *minimal sequent calculus*, such that

- the set of formulas is SEQ,
- the set LA of logical axioms is $\{(X, x) : x \in X\}$,
- there is a finitary rule:

$$\frac{(X, \alpha)}{(X \cup Y, \alpha)}$$

  and an infinitary rule:

$$\frac{\{(X, \beta) : \beta \in Z\}, (X \cup Z, \alpha)}{(X, \alpha)}.$$

A theory of such a system is a class $\vdash$ of sequents such that

- $\vdash \supseteq LA$
- $(X, \alpha) \in \vdash \Rightarrow (X \cup Y, \alpha) \in \vdash$

- $\{(X,\beta) : \beta \in Z\} \subseteq \vdash$ and $(X \cup Z, \alpha) \in \vdash \Rightarrow (X, \alpha) \in \vdash$.

The proof of the following proposition is evident.

Proposition 3.1. *The class of theories of the minimal sequent calculus coincides with the class of consequence relations.*

In [1] and [9], one proposes the notion of canonical extension of an inferential apparatus into a fuzzy deduction apparatus. The idea is that in the canonical extension the set of axioms remains the same while every n-ary inference rule $r' : F^n \to F$ is extended into a fuzzy inference rule by adding the function $r'' : U^n \to U$ defined by setting $r''(\lambda_1, ..., \lambda_n) = \lambda_1 \wedge ... \wedge \lambda_n$. The interpretation is that if $\alpha_1, ..., \alpha_n$ were proved with degree $\lambda_1, ..., \lambda_n$ then the conclusion $r'(\alpha_1, ..., \alpha_n)$ is proved with degree $\lambda_1 \wedge ... \wedge \lambda_n$. Such a definition is extended to the case a rule admits infinitely many premises in an obvious way.

Now, the 'canonical extension' of the minimal sequent calculus, we denote by Σ^*, is defined by assuming that:

- $la(X, x) = 1$ if $x \in X$, and $la(X, x) = 0$, otherwise
- there is a finitary rule

$$\frac{(X,\alpha)}{(X \cup Y, \alpha)} \ ; \ \frac{\lambda}{\lambda}$$

- there is an infinitary rule

$$\frac{\{(X,\beta) : \beta \in Z\}, (X \cup Z, \alpha)}{(X, \alpha)} \ ; \ \frac{S, \lambda}{\inf(s) \wedge \lambda}.$$

The proof of the following theorem is trivial.

Theorem 3.2. *The class of graded consequence relations coincides with the class of theories of the canonical extension Σ^* of the minimal sequent calculus Σ.*

As an immediate consequence, we obtain the following corollary.

Corollary 3.2. *Given any graded sf-relation g, we can consider the graded consequence $D(g)$ generated by g, i.e. the least graded consequence extending g.*

4 Finite sequent calculus and compact graded consequences

The notion of compactness is a basic one in any logic. Indeed such a notion emphasizes that the inferential process is finite in nature. In connection with the graded consequence theory, we propose the following definition where, given a set X, we indicate by $P_f(X)$ the class of finite subsets of X.

Definition 4.1. *We say that a sf-relation* $\vdash$ *is* compact *if*

$$X \vdash \alpha \Longleftrightarrow \text{ there exists } X_f \in P_f(X) \text{ such that } X_f \vdash \alpha.$$

We say that a graded sf-relation g *is* compact *if*

$$g(X \vdash \alpha) = \sup\{g(X_f \vdash \alpha) : X_f \in P_f(X)\}. \tag{5}$$

We can characterize the compact graded consequences as follows.

Proposition 4.1. *Let* g *be a graded sf-relation, then* g *is a compact graded consequence if and only if* g *is compact and*

(i) $g(X \vdash \alpha) = 1$ *for every* $\alpha \in X$,
(ii) $g(X \cup Y \vdash \alpha) \geq g(X \vdash \alpha)$,
(iii) $g(X \vdash \alpha) \geq g(X \vdash z) \wedge g(X \cup \{z\} \vdash \alpha)$.

Proof. It is evident that if g is a compact graded consequence, then (i), (ii) and (iii) are satisfied. Conversely, assume that g is compact and that (i), (ii) and (iii) are satisfied. Then, first we demonstrate that, for every finite set Z_f,

$$g(X \vdash \alpha) \geq (\inf\{g(X \vdash z) : z \in Z_f\}) \wedge g(X \cup Z_f \vdash \alpha).$$

We put $Z_f = \{z_1, ..., z_n\}$ and we prove this by induction on n. Indeed, in the case $n = 1$ such an inequality coincides with (iii). Assume that $n \neq 1$, then, by induction hypothesis and by (iii),

$$\begin{aligned} g(X \vdash \alpha) &\geq (\inf\{g(X \vdash z) : z \in \{z_1, ..., z_{n-1}\}\}) \wedge g(X \cup \{z_1, ..., z_{n-1}\} \vdash \alpha) \\ &\geq (\inf\{g(X \vdash z) : z \in \{z_1, ..., z_{n-1}\}\}) \wedge g(X \cup \{z_1, ..., z_{n-1}\} \vdash z_n) \\ &\wedge g(X \cup \{z_1, ..., z_{n-1}, z_n\} \vdash \alpha) \\ &\geq (\inf\{g(X \vdash z) : z \in \{z_1, ..., z_{n-1}\}\}) \wedge g(X \vdash z_n) \wedge g(X \cup Z_f \vdash \alpha) \\ &= (\inf\{g(X \vdash z) : z \in Z_f\}) \wedge g(X \cup Z_f \vdash \alpha). \end{aligned}$$

Let Z be any set of formulas, and consider the supremum with respect to the finite subsets of Z. Then,

$$\begin{aligned} g(X \vdash \alpha) &\geq \sup\{\inf(\{g(X \vdash z) : z \in Z_f\}) \wedge g(X \cup Z_f \vdash \alpha) : Z_f \in P_f(Z)\} \\ &\geq \sup\{(\inf\{g(X \vdash z) : z \in Z\}) \wedge g(X \cup Z_f \vdash \alpha) : Z_f \in P_f(Z)\} \\ &= (\inf\{g(X \vdash z) : z \in Z\}) \wedge (\sup\{g(X \cup Z_f \vdash \alpha) : Z_f \in P_f(Z)\} \\ &= (\inf\{g(X \vdash z) : z \in Z\}) \wedge g(X \cup Z \vdash \alpha) \end{aligned}$$

where the last equality is justified by the compactness of g ◻

We call *finite sequent* any sequent (X, α) in which X is finite and we denote by SEQ_f the set of finite sequents. If $\vdash$ is compact, then $\vdash$ is completely determined by its intersection with SEQ_f. If $\vdash$ is a subset of SEQ_f, we define $\underline{\vdash}$ by setting

$$\underline{\vdash} = \{(X, \alpha) \in SEQ : \text{there is a finite subset} X_f \subseteq X, X_f \vdash \alpha\}.$$

In such a case we say that $\underline{\vdash}$ is the *compact extension* of $\vdash$.

Likewise, a compact graded *sf*-relation g is completely defined by its restriction to SEQ_f. Let h be a fuzzy subset of SEQ_f and set

$$\underline{h}(X,\alpha) = \sup\{h(X_f,\alpha) : X_f \text{ is a finite subset of } X\}.$$

Then $\underline{h}$ is a compact graded *sf*-relation we call *the compact extension of* h.

Proposition 4.2. *g is a compact graded consequence iff g is the compact extension of a graded sf-relation $h : SEQ_f \to U$ satisfying*

(j) $h(X,\alpha) = 1$ *for every* $\alpha \in X$,
(jj) $h(X \cup Y,\alpha) \geq h(X,\alpha)$,
(jjj) $h(X,\alpha) \geq h(X,\beta) \wedge h(X \cup \{\beta\},\alpha)$.

Proof. If g is a compact graded *sf*-relation, then it is obvious that its restriction h to SEQ_f satisfies (j), (jj) and (jjj). Conversely, let g be the compact extension of a graded *sf*-relation h satisfying (j), (jj) and (jjj). Then, by Proposition 4.1, to prove g is a graded consequence it is sufficient to prove that g satisfies (i) and (ii) and

$$g(X,\alpha) \geq g(X,\beta) \wedge g(X \cup \{\beta\},\alpha). \tag{6}$$

Now, (i) and (ii) are trivial. In order to prove (6) observe that $g(X,\alpha) = \sup\{h(X_f,\alpha) : X_f \in P_f(X)\} \geq \sup\{h(X_f,\beta) \wedge h(X_f \cup \{\beta\},\alpha) : X_f \in P_f(X)\}$. On the other hand,

$$\begin{aligned} &g(X,\beta) \wedge g(X \cup \{\beta\},\alpha) \\ &\qquad = (\sup\{h(X_1,\beta)) : X_1 \in P_f(X)\}) \wedge (\sup\{h(X_2 \cup \{\beta\},\alpha) : X_2 \in P_f(X)\}) \\ &\qquad = \sup\{h(X_1,\beta) \wedge h(X_2 \cup \beta\},\alpha) : X_1, X_2 \in P_f(X)\}. \end{aligned}$$

Now, observe that, if $X_f = X_1 \cup X_2$, then $h(X_f,\beta) \wedge h(X_f \cup \beta,\alpha) \geq h(X_1,\beta) \wedge h(X_2 \cup \{\beta\},\alpha)$ and therefore

$$\begin{aligned} g(X,\alpha) &\geq \sup\{h(X_f,\beta) \wedge h(X_f \cup \{\beta\},\alpha) : X_f \in P_f(X)\} \\ &\geq \sup\{h(X_1,\beta) \wedge h(X_2 \cup \beta\},\alpha) : X_1, X_2 \in P_f(X)\} \\ &= g(X,\beta) \wedge g(X \cup \{\beta\},\alpha). \end{aligned}$$

Then, we can conclude that (6) holds true. □

Such a proposition enables us to eliminate the infinitary inference rule in defining the canonical extension of the minimal sequent calculus. In fact, let $\Sigma_f = (LA, IR)$ be the H-system such that

- SEQ_f is the set of formulas,
- the set LA of logical axioms is $\{(X,x) \in SEQ_f : x \in X\}$,
- there are the following rules:

$$\frac{(Y,\alpha)}{(X \cup Y,\alpha)} \quad ; \quad \frac{(X,\beta),(X \cup \{\beta\},\alpha)}{(X,\alpha)}.$$

We call *minimal finite-sequent calculus* such a system. Then, we have the following theorem whose proof is immediate.

Theorem 4.1. *Let Σ_f^* be the canonical extension of Σ_f. Then, g is a compact graded consequence iff g is the compact extension of a theory of Σ_f^*.*

5 A connection with the stratified closure operators

Let L be a complete lattice. Then a *closure operator* in L is an operator $J : L \to L$ satisfying:

(i) $x \leq y \Rightarrow J(x) \leq J(y)$ (order-preserving),
(ii) $x \leq J(x)$ (inclusion),
(iii) $J(J(x)) = J(x)$ (idempotence).

An order-preserving operator J is called *continuous* provided that, for every directed class C of elements of L

$$J(\lim C) = \lim J(C).$$

Given a set S, we are interested to the case in which L is the lattice of all the subsets of a given set S. In such a case the continuity of an operator coincides with the compactness of such an operator. As it is well known, the deduction operator of a crisp logic is a compact closure operator in the lattice $P(F)$. Also, we are interested to the case L is the lattice of all the fuzzy subset of S. We call *fuzzy operator on S* an operator in such a lattice. In this case the continuity is again a kind of compactness. Indeed, denote by U_f^S the class of the finite fuzzy subsets of S with rational values and write $s_1 \ll s_2$ provided that $s_1(x) < s_2(x)$ for every x in $Supp(s_1)$. Then J is continuous if and only if

$$J(v) = \bigcup \{J(s_f) : s_f \ll s\}.$$

It is easy to prove that the deduction operator in a fuzzy logic is a continuous closure operator in the lattice U^F. This suggests to propose an abstract approach to fuzzy logic intended as a theory of continuous closure operators in U^F.

This section is devoted to compare the notion of fuzzy closure operator with the one of graded consequence. At first we observe that every operator J in $P(F)$ is associated with a *sf*-relation $\vdash$ by setting

$$X \vdash \alpha \Leftrightarrow \alpha \in J(X).$$

Moreover, every *sf*-relation $\vdash$ is associated with an operator J defined by setting

$$J(X) = \{\alpha \in F : X \vdash \alpha\}.$$

The proof of the following proposition is matter of routine.

Proposition 5.1. *Given a closure operator J, its associated sf-relation is a consequence relation $\vdash$. If J is compact, then $\vdash$ is compact. Conversely, given a consequence relation $\vdash$, its associated operator J is a closure operator. If $\vdash$ is compact, then J is compact.*

The question arises whether we can extend such a connection to the graded consequence relations and the fuzzy closure operators in F. Notice that while the graded consequences operate on sets of formulas, the fuzzy operators operate on fuzzy subsets of formulas. Now, if $J : U^F \to U^F$ is a fuzzy closure operator, then we can define a graded *sf*-relation g by setting

$$g(X \vdash \alpha) = J(X)(\alpha) \tag{7}$$

Also, if g is a graded consequence and $\vdash_\lambda$ denotes its $\lambda-$cut, then we can define a fuzzy operator D by setting

$$D(s)(\alpha) = \sup\{\lambda : C(s, \lambda) \vdash_\lambda \alpha\} \tag{8}$$

Regarding Definition (7), we have the following proposition.

Proposition 5.2. *Let $J : U^F \to U^F$ be a fuzzy closure operator and define the graded sf-relation g by (6). Then, by referring to Definition 3.1, g satisfies (i) and (ii) but not (iii) in general. Therefore g is not a graded consequence.*

Proof. A straightforward verification proves the first part of the proposition. In the following example (iii) is not satisfied (M. K. Chakraborty, personal communication). Let $F = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ and let s_1 and s_2 be the two fuzzy subsets of F defined by setting $s_1(\alpha_1) = s_1(\alpha_3) = 1, s_1(\alpha_2) = 0.7, s_1(\alpha_4) = 0.8$ and $s_2(\alpha_1) = s_2(\alpha_3) = s_2(\alpha_4) = 1, s_2(\alpha_2) = 0.9$.

Define J by setting, for every fuzzy subset s and $\alpha \in F$,

$$J(s)(\alpha) = \inf\{s_i(\alpha) : s_i \supseteq s\}.$$

Take $X = \{\alpha_1, \alpha_3\}$ and $Z = \{\alpha_4\}$. Then, a simple calculation gives $J(X)(\alpha_2) = 0.7, J(X \cup Z)(\alpha_2) = 0.9, J(X)(\alpha_4) = 0.8$. So, if g is the graded *sf*-relation associated with J, $g(X \vdash \alpha_2) = 0.7, \inf\{g(X \vdash z) : z \in Z\} = 0.8,$ and $g(X \cup Z \vdash \alpha_2) = 0.9$. Hence, $g(X \vdash \alpha_2) < (\inf\{g(X \vdash z) : z \in Z\}) \wedge g(X \cup Z \vdash \alpha_2)$. This demonstrates that (iii) is not satisfied. □

To obtain that g is a graded consequence we have to refer to a particular class of closure operators: the well stratified closure operators. To introduce such a notion, imagine a class of different deduction tools each with a related degree of validity. We can represent such a state of affairs assuming that, for every $\lambda \in U$, a crisp deduction operator D_λ is defined. This means that, for every set X, we interpret $D_\lambda(X)$ as the set of formulas that we can derive from X by argumentations which are 'reliable' with degree λ. So, the deduction apparatus is 'stratified', in a sense. More generally, it is possible that both the available information and the deduction apparatus are stratified. In this case, if we represent the stratified information by a fuzzy set $v : F \to [0, 1]$ and

$\alpha \in D_\lambda(C(v,\lambda))$, we say that α *is a consequence of* v *at least with degree* λ. This means that λ is a lower bound constraint on the truth value of α. Obviously, given a formula α, we have to search for the best lower-constraint we are able to get. This leads to the following definition.

Definition 5.1. *We say that a family* $(D_\lambda)_{\lambda \in U}$ *of closure operators is a* continuous chain *provided that* $(D_\lambda(X))_{\lambda \in U}$ *is a continuous chain for every subset* X. *In such a case we associate* $(D_\lambda)_{\lambda \in U}$ *with the fuzzy operator* D *defined by setting, for every* $s \in U^F$ *and* $\alpha \in F$,

$$D(s)(\alpha) = \sup\{\lambda \in U : \alpha \in D_\lambda(C(s,\lambda))\}. \tag{9}$$

The proof of the following theorem is in [9].

Theorem 5.1. *Let* $(D_\lambda)_{\lambda \in U}$ *be a continuous family of closure operators and let* D *be the associated operator. Then* D *is a fuzzy closure operator (we call well-stratified).*

Theorem 5.2. *A graded sf-relation* $g : P(F) \times F \to U$ *is a graded consequence iff a well-stratified closure operator* D *exists such that*

$$g(X \vdash \alpha) = D(X)(\alpha) \tag{10}$$

for every X *subset of* F *and* $\alpha \in F$.

Proof. Let g be a graded consequence and, for every $\lambda \in U$, let D_λ be the closure operator associated with the consequence relation $\vdash_\lambda = C(g,\lambda)$, i.e.

$$D_\lambda(X) = \{\alpha \in F : X \vdash_\lambda \alpha\}.$$

Moreover, denote by D the closure operator associated with $(D_\lambda)_{\lambda \in U}$. Then $g(X \vdash \alpha) = \sup\{\lambda \in U : g(X \vdash \alpha) \geq \lambda\} = \sup\{\lambda \in U : \alpha \in D_\lambda(X)\} = D(X)(\alpha)$. So, we must prove only that $(D_\lambda)_{\lambda \in U}$ is a continuous chain. Let X be a set of formulas. Then, trivially, $D_0(X) = F$. Furthermore, if $\mu \in U$, then $\alpha \in D_\mu(X) \Leftrightarrow g(X \vdash \alpha) \geq \mu \Leftrightarrow g(X \vdash \alpha) \geq \lambda$ for every $\lambda < \mu \Leftrightarrow \alpha \in \cup_{\lambda<\mu} D_\lambda(X)$.

Conversely, let (F, D) be the fuzzy deduction system associated with a given continuous chain $(D_\lambda)_{\lambda \in U}$ of deduction systems and, for every $\lambda \in U$, denote by $\vdash_\lambda$ the consequence relation associated with D_λ , that is $\vdash_\lambda = \{(X,x) : x \in D_\lambda(X)\}$. We claim that $(\vdash_\lambda)_{\lambda \in U}$ is a continuous family. Indeed, $\vdash_0 = SEQ$ and $X \vdash_\mu \alpha \Leftrightarrow \alpha \in D_\mu(X) \Leftrightarrow \alpha \in D_\lambda(X)$ for every $\lambda < \mu \Leftrightarrow X \vdash_\lambda \alpha$ for every $\lambda < \mu$. Thus, by Theorem 3.1 the graded *sf*-relation g defined by (10) is a graded consequence. □

We observe in an explicit way that this theorem entails that any graded consequence g can be represented as follows

$$g(X \vdash \alpha) = \sup\{\lambda \in U : \alpha \in D_\lambda(X)\}$$

where $(D_\lambda)_{\lambda \in U}$ is a continuous family of closure operators.

Another question is to associate any graded consequence g with a suitable fuzzy closure operator D. To this purpose we can set Now, it is possible to so this purpose, lso, such a theorem suggests a natural way to extend a graded consequence relation $g : P(F) \times F \to U$ in a fuzzy relation $\tilde{g} : U^F \times F \to U$. Indeed, it is sufficient to consider the stratified deduction operator D associated with g and to set

$$\tilde{g}(s \vdash \alpha) = D(s)(\alpha). \tag{11}$$

Equivalently, we can set

$$\tilde{g}(s \vdash \alpha) = \sup\{\lambda \in U : C(s, \lambda) \vdash_\lambda \alpha\} \tag{12}$$

for every $s \in U^F$ and $\alpha \in F$. This suggests to examine the possibility of extending the theory of the graded consequence by calling *fuzzy sequent* any pair (s, α) and by defining the graded consequence relations as suitable fuzzy relations between fuzzy subsets and formulas.

6 Introducing effectiveness

The effectiveness of the deduction operator is an essential feature for any logic. In particular, in classical logic this effectiveness is expressed by the fact that if X is a decidable set of axioms then the related set $D(X)$ of theorems is effectively enumerable. In addition, the passage from an algorithm for (the characteristic function of) X to an algorithm for (the effective enumeration of) $D(X)$ is obtained in an "uniform way". Now the better way to represent this is by the notion of "enumeration operator" or "computable operator" we express in terms of sequents (see Roger's book [15])).

Definition 6.1. *An operator $H : P(F) \to P(F)$ is* computable *if a recursively enumerable subset W of SEQ_f exists such that*

$$H(X) = \{\alpha \in F : \textit{there is } (X_f, \alpha) \in W, X_f \subseteq X\}. \tag{13}$$

Then H is computable provided that a computable map $h : N \to SEQ_f$ exists such that, for any $X \in P(F)$,

$$H(X) = \{\alpha \in F : \text{there is } n \in N \text{ such that } h(n) = (X_f, \alpha), X_f \subseteq X\}. \tag{14}$$

It is easy to verify that if X is effectively enumerable, in particular if X is decidable, then $H(X)$ is effectively enumerable. The following proposition, whose proof is matter of routine, gives an alternative definition.

Proposition 6.1. *An operator $H : P(F) \to P(F)$ is computable if and only if it is compact and the relation*

$$W_H = \{(X_f, \alpha) \in SEQ_f : \alpha \in H(X_f)\}$$

is a recursively enumerable sf-relation.

To extend such a notion to the *sf*-relations we can use the connection between sf-relations and operators established in Section 5.

Definition 6.2. *A sf-relation* $\vdash$ *is* computable *if its associated operator is computable.*

Then $\vdash$ is computable if a computable map $h : N \to SEQ_f$ exists such that, for any $X \in P(F)$,

$$X \vdash \alpha \Leftrightarrow \text{there is } n \in N \text{ such that } h(n) = (X_f, \alpha) \text{ and } X_f \subseteq X. \tag{15}$$

We can interpret this by assuming that there is an effective coding of the set of the proofs in a logic and that for every proof n, $h(n) = (X_f, \alpha)$ means that n proves α and that its set of hypotheses is X_f. Then (15) says that $X \vdash \alpha$ holds true if and only if a proof of α exists whose set of hypotheses are in X.

Proposition 6.2. *A sf-relation* $\vdash$ *is computable if and only if* $\vdash$ *is the compact extension of an effectively enumerable sf-relation* $W \subseteq SEQ_f$ *if and only if* $\vdash$ *is compact and its restriction to the finite sets of formulas is effectively enumerable.*

Proof. Assume that $\vdash$ is computable and let h be as in (15). Then, since the subset $\{(n, X) : "h(n) = (X_f, \alpha) \text{ and } X_f \subseteq X"\}$ of $N \times P_f(F)$ is decidable, the restriction of $\vdash$ to finite sets of formulas is recursivery enumerable. The converse implication is trivial. In a similar way one proves the remaining part of the proposition. □

To extend Definitions 6.1 and 6.2 to the fuzzy framework, at first we give the definition of effectively enumerable fuzzy subset of a set S (see [2], [9], [10]). Obviously, we refer to a set S for which there is a coding in N. By U_Q we denote the set of rational numbers in U.

Definition 6.3. *A fuzzy subset* $s : S \to U$ *of a set* S *is* effectively enumerable *if an effectively computable function* $h : S \times N \to U_Q$ *exists which is increasing with respect to the second variable and such that, for every* $x \in S$,

$$s(x) = \lim_{n\to\infty} h(x, n).$$

Then a fuzzy subset s is effectively enumerable provided that, for every $x \in S$, we can calculate $s(x)$ by an effective approximation process “from below”.

Definition 6.4. *A fuzzy subset* $s : S \to U$ *of a set* S *is* decidable *provided that both* s *and its complement* $-s$ *are effectively enumerable.*

Equivalently, s is decidable provided that there are two computable maps $h : S \times N \to U_Q$ and $k : S \times N \to U_Q$, order-preserving and order-reversing with respect to the second variable, respectively, and such that for every $x \in S$, $\lim_{n\to\infty} h(x, n) = s(x) = \lim_{n\to\infty} k(x, n)$. In other words, we can calculate $s(x)$ by an effective approximation process “from below” and an effective approximation process “from above”. It is easy to see that by substituting U with the

Boolean algebra $\{0,1\}$ we obtain the usual definitions of effectively enumerable and decidable subset.

Set $\underline{SEQ}_f = U_f^S \times F$. Then, since an element of such a set is defined by a finite set of formulas and rational numbers, there is no difficulty to define a coding for $\underline{SEQ}_f$ and therefore the notion of effectively enumerable fuzzy subset of $\underline{SEQ}_f$.

Definition 6.5. *We say that a fuzzy operator $H : U^F \to U^F$ is* computable *if an effectively enumerable fuzzy subset $w : \underline{SEQ}_f \to U$ of $\underline{SEQ}_f$ exists such that, for every $\alpha \in F$,*

$$H(s)(\alpha) = \sup\{w(s_f \vdash \alpha) : s_f \ll s\}. \tag{16}$$

Then H is computable if a recursive map $k : \underline{SEQ}_f \times N \to U_Q$ exists such that, for any $\alpha \in F$ and $s \in U^F$,

$$H(s)(\alpha) = \sup\{k(s_f, \alpha, n) : n \in N \text{ and } s_f \ll s\}. \tag{17}$$

The following theorem gives some support to the just given definition.

Theorem 6.1. *Let $H : U^F \to U^F$ be computable. Then H is continuous. Moreover, if s is effectively enumerable, then $H(s)$ is effectively enumerable. In particular*

$$s \text{ decidable } \Rightarrow H(s) \text{ effectively enumerable }.$$

Proof. Let $(s_i)_{i\in I}$ be a directed family of fuzzy subsets and observe that if $s \in U_f^S$ and $s \ll \cup_{i\in I} s_i$, then there is $i \in I$ such that $s \subseteq s_i$. Then

$H(\cup_{i\in I} s_i)(\alpha) = \sup\{w(s_f \vdash \alpha) : s_f \ll \cup_{i\in I} s_i\} = \sup\{w(s_f \vdash \alpha) : s_f \ll s_i, i \in I\} = \sup_{i\in I} H(s_i)(\alpha)$.

Assume that s is recursively enumerable and let h be as in Definition 6.3. Moreover, let $k : \underline{SEQ}_f \times N \to U_Q$ be as in (17) and, for every $n \in N$, let s_n be the fuzzy subset of formulas defined by setting $s_n(\alpha) = h(\alpha, n)$ for any $\alpha \in F$. Then, since H is continuous and $(s_n)_{n\in N}$ is a directed family, we have $H(s)(\alpha) = H(\cup_{n\in N} s_n)(\alpha) = \sup_{n\in N} H(s_n)(\alpha)$. By observing that $H(s_n)(\alpha) = \sup\{w(s_f \vdash \alpha) : s_f \ll s_n\}$, we get $H(s)(\alpha) = \sup\{k(s_f, \alpha, m) : s_f \ll s_n, n, m \in N\}$. Define r by setting $r(\alpha, i, n, m) = k(s_f, \alpha, m)$ if i is the code number of s_f and $s_f \ll s_n$ and otherwise, $r(\alpha, i, n, m) = 0$. Then r is a recursive map and

$$H(s)(\alpha) = \sup_{i\in N} \sup_{n\in N} \sup_{m\in N} r(\alpha, i, n, m).$$

It is easy to see that this entails that $H(s)$ is recursively enumerable. □

Theorem 6.2. *Given $H : U^F \to U^F$, define $w_H : \underline{SEQ}_f \to U$ by setting, for any $\alpha \in F$ and $s_f \in U_f^S$,*

$$w_H(s_f \vdash \alpha) = H(s_f)(\alpha). \tag{18}$$

Then H is computable if and only if H is continuous and w_H is a recursively enumerable fuzzy relation.

Proof. Let H be computable and let k be as in (17). Then $w_H(s_f \vdash \alpha) = H(s_f)(\alpha) = \sup\{k(s'_f, \alpha, m) : s'_f \ll s_f, m \in N\}$. Define the function h by setting $h(s'_f, \alpha, s_f, m) = k(s'_f, \alpha, m)$ if $s'_f \ll s_f$ and otherwise, $h(s'_f, \alpha, s_f, m) = 0$. Then $w_H(s_f \vdash \alpha) = \sup\{h(s'_f, \alpha, s_f, m) : s'_f \in U_f^S$ and $m \in N\}$, and this proves that w_H is recursively enumerable.

Conversely, if w_H is recursively enumerable and H is continuous, then $H(s)(\alpha) = \sup\{H(s_f)(\alpha) : s_f \ll s\} = \sup\{w(s_f \vdash \alpha) : s_f \ll s\}$. This proves that H is computable. $\square$

To attempt for a definition of computable graded sf-relation a first idea is to use the connection between these relations and the computable fuzzy operators established in Section 5.

Definition 6.6. *Let g be a graded sf-relation, then we say that g is* computable *if g is the graded sf-relation associated with a computable operator.*

Then g is computable if there is a computable map $k : \underline{SEQ_f} \times N \to U_Q$ increasing with respect to n such that

$$g(X \vdash \alpha) = \sup\{k(s_f, \alpha, n) : n \in N, s_f \ll X\}.$$

Such a definition is rather obscure since it involves the finite fuzzy subsets s_f to define a relation on classical subsets. Perhaps we can propose the following one arising from the definition of the compactness of a graded sf-relations.

Definition 6.7. *Let g be a graded sf-relation, then we say that g is* computable *if g is compact and its restriction to SEQ_f is effectively enumerable.*

7 Conclusions

We conclude this paper by emphasizing that all the questions on graded consequences sketched in this paper requires further investigations. As an example, the connection with the canonical extension has to be extended in order to admit any triangular norm. Regarding the connections between graded consequences and fuzzy operators, it would be interesting to prove that such a connection preserves the compactness. Also, should be useful to prove the equivalence between Definition 9 and Definition 10.

References

1. Biacino L., Gerla G., [1996], An extension principle for closure operators, J. of Math. Anal. Appl., 198, 1-24.
2. Biacino L., Gerla G. [2002], Fuzzy logic, continuity and effectiveness, Archive for Mathematical Logic, 41, 643-667.
3. Chakraborty M.K. [1988], Use of fuzzy set theory in introducing graded consequence in multiple valued logic, in Fuzzy Logic in Knowledge-Based Systems, Decision and Control, M.M. Gupta and T.Yamakawa (eds), North-Holland, Amsterdam, 247-257.

4. Chakraborty M.K. [1994], Graded consequence: further studies, J. of Applied non-classical Logic, 5, 227-237.
5. Chakraborty M.K., Dutta S. [2009], Some Proof Theoretic Results Depending on Context from the Perspective of Graded Consequence, 2009, 144-151.
6. Chakraborty M.K., Dutta S.,Graded Consequence Revisited, Fuzzy Sets and Systems, 161 (2010), 1885-1905.
7. Dubois D., Lang J., Prade H. [1994], Possibilistic logic, In: D.Gabbay, C.Hogger and J.Robinson (eds.), Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 3. Clarendon Press.
8. Gerla G. [1996], Graded consequence relations and closure operators, J. of Applied non-classical Logic, 6, 369-379.
9. Gerla G. [2001], Fuzzy Logic: Mathematical Tools for Approximate Reasoning. Kluwer Academic Publishers.
10. Gerla G.[2006], Effectiveness and Multivalued Logics, Journal of Symbolic Logic, 71, 137-162.
11. Gottwald S. [1993], Fuzzy Sets and Fuzzy Logic, Wiesbaden: Vieweg.
12. Hájek P. [1998], Metamathematics of Fuzzy Logic, Kluwer Academic Publishers, Dordrecht.
13. Novák V., Perfilieva I, Mockor J. [1999], Mathematical Principles of Fuzzy Logic, Kluwer Academic Publishers, Dordrecht.
14. Pavelka J. [1979], On fuzzy logic I: Many-valued rules of inference, Zeitschr. f. math. Logik und Grundlagen d. Math., 25, 45-52.
15. Rogers H. [1967], Theory of Recursive Functions and Effective Computability, McGraw-Hill, New York.
16. Turunen E. [1999], Mathematics Behind Fuzzy Logic, Physica-Verlag.
17. Zadeh L.A. [1965], Fuzzy Sets, Information and Control, 12, 338-353.

Cristina Coppola, Giangiacomo Gerla
Department of Mathematics,
University of Salerno, Italy.
ccoppola@unisa.it, ggerla@unisa.it

Three-valued logics and Knowledge Representation: Pragmatic Issues

Didier Dubois

1 Introduction

There has been in the last thirty years or so a considerable surge of interest in the issue of uncertainty handling in information processing, decision theory and artificial intelligence. However, the tradition of logic did not so much focus on this issue in the first half of XXth century. Indeed, it is well-known that since the divorce between probability and logic more than 100 years ago, the latter has been considered more in the scope of the foundations of mathematics and more recently computer sciences, while probability theory became the major tool for representing belief and uncertainty after the pioneering works of De Finetti and Ramsey (see [34] for a collection of seminal papers). In the 1960's the raise of modal logics gave an opportunity to develop logics of knowledge and belief [30], as initiated by Hintikka [31], originally without clear connections with probability. Since then, artificial intelligence and decision theory have motivated the need for moving away from the strict Bayesian probability tradition, by introducing explicit representations of incomplete information. Various kinds of concerns led to this trend

- On the one hand, there was the necessity of distinguishing between not believing and believing that not in expert systems with certainty factors (something modal logic easily grasps, albeit in a symbolic approach): it ended up with the development of evidence theory [38].
- The representation of gradual linguistic information supplied by humans. Initiated by Zadeh, this topic led to possibility theory [41].
- The limited expressiveness of relational tables for representing sets of objects in databases, due to indiscernibility phenomenon: this is the theory of rough sets [36].

Another logic trend, initiated by Łukasiewicz in the 1920's, is multiple logics [35]. However, while the algebraic setting developed in connection with the multiple-valued logic trend turned out to be very fruitful, the meaning of truth-degrees seemed to be problematic from the start. In particular, a number of authors have tried to attach an epistemic flavor to such degrees, starting with Łukasiewicz himself, for whom the third truth-value different from *true* or *false*, explicitly referred to the idea of *possible* (in connection with future contingents), which is a modality. And indeed there have been attempts to capture three-valued logics by modal logics, as well as attempts at modeling incomplete information problems by three-valued logics. This situation has led to some confusion between degrees of truth and degrees of belief in many-valued logical formalisms. However, the situation starts to be clarified, due to a number of works due to Hájek [28], Dubois and Prade [18, 15], and others, trying to disentangle the two issues of gradual truth and gradual belief in the logical setting.

Interestingly, the works of Chakraborty also deal with interactions between many-valued logics and incomplete information. On the one hand, he made efforts at capturing degrees of entailment [10, 11], and it is clear that viewing a set of propositions as a set of beliefs, there is a close connection between entailment and belief, since believed propositions are those that can be entailed by the set of propositions. On the other hand, Chakraborty and his students investigated logical foundations of rough sets, and retrieve the three-valued logic originally proposed by Łukasiewicz [4, 3, 37].

This short note discusses the issue of semantics of three-valued logics in relation to the handling of incomplete, conditional or gradual information, relying on personal discussions with Chakraborty. A central thesis is that rather than trying to preserve a simple syntax (like the one of propositional logics) and overloading the semantics by epistemic notions, it is better to use two levels of syntax, one that speaks of properties of the real world, and another that handles epistemic notions, in the tradition of epistemic modal logics, but without resorting to relational semantics à la Kripke, whose target is to handle nested modalities, which is of no concern in this paper. The paper also relies on some recent works by the author, and insists on lines for future research.

2 Semantic issues on three-valued logics

The role of Boolean logic in the philosophical and mathematical tradition has been overwhelming, and for many, the notion of truth cannot be gradual. This state of facts is particularly patent in studies about vagueness, where the use of truth-degrees, taken for granted in fuzzy logic, is severely criticized by philosophers like Fine [23] or Williamson [40]. There is a dispute between tenants of gradual truth approaches to vagueness, and those who interpret this phenomenon in terms of epistemic uncertainty, truth-value gaps, or the possibility of several models. We have argued elsewhere that the latter views can be reconciled with gradual truth [16].

In fact, while Boolean truth is hard to dispute in a purely philosophical perspective, taking the point of view of computer and information sciences opens the door to natural interpretations of more than two truth-values. And first of all, the role of logic is then to represent information. This means a language is needed. If we assume simplicity, objects are then described by means of properties, which can be modelled by atomic propositions. If an object possesses such a property, the corresponding atomic proposition is true. If the property is Boolean, it can be attached to a variable taking values on the set $\{T, F\}$, where T stand for *true* and F for *false*. In this situation it is clear that

- *true* is just a value in the domain of the attribute describing the property, not the absolute metaphysical notion traditional philosophy speaks about.
- The traditional notion of truth comes down to evaluating the conformity of a statement with respect to the real state of facts. But, in the information representation perspective, only pieces of information can be compared. The real state of facts is not accessible, only its description can be and the latter requires a language whose precision is limited. The set of interpretations of a language only describes equivalence classes of states of facts (and this is precisely the point made by rough sets).
- As a consequence, the choice of a truth-set, viewed as an attribute domain is a matter of convention, since the choice of the language is ours, as argued by De Finetti [14]. One may decide whether a property like *tall* will be modelled by a Boolean predicate, or if we need a more refined truth-scale that mirrors the age scale in a more faithful way.

Of course we can always describe a gradual property with a finite number n of truth-values by a set of k Boolean variables with $2^{k-1} \leq n \leq 2^k$, plus axioms expressing exhaustiveness and mutual exclusions of gradual values in the scale, but this approach looks tedious.

If we consider the choice of using three truth-values instead of two, it means we would work with a three-valued chain $(\{F, M, T\}, \geq)$ where $F < M < T$. While this setting is slightly more elaborated that the usual Boolean setting, one may admit that it sounds almost as basic as the latter, even if it allows for more connectives (especially binary) than Boolean logics. Prior to defining a three-valued algebraic framework, the question is then to find natural understandings of the third truth value M (for "middle") in such a way that propositions in the given language could be declared to be M in some state of affairs.

In the literature one can find several possible understandings of the third truth-value M. Here are four examples:

- **undefined** : this view was put forward by Kleene [32], viewing logical formulas as recursively defined functions, and moving to partial recursive functions. The idea is that if one component of the recursive function is undefined in some interpretation the whole function is undefined. This interpretation of M seems to be akin to the motivation of Bochvar logic, where M refers to meaninglessness.
- **unknown** : this view is perhaps the most usual proposal and was put forward by Kleene [32] in his book, namely (p. 64):

  > We further conclude [...] that, for the definitions of partial recursive operations, T, F, U must be susceptible of another meaning besides (i) 'true', 'false', 'undefined', namely (ii) 'true', 'false', 'unknown' (or value immaterial). Here, 'unknown' is a category into which we regard any proposition as falling, whose value we either do not know or choose for the moment to disregard; and it does not exclude the two other possibilities 'true' and 'false'.

  This approach has been considered as a natural way of handling incomplete information in logic and has been developed further in partial logic [8], for instance. More recently it has been found natural to apply it to formal concept analysis to handle contexts with missing information [24]. The intuition of Lukasiewicz that M means *possible* can be thought of as akin to the idea of not knowing, as it makes reference to future contingent (the idea of *not known yet*), even though Lukasiewicz logic differs from Kleene's.
- **irrelevant**: in this case, if a proposition takes this truth-value, it should not affect the truth-value of formulas containing this proposition as a component. The reason is that this proposition becomes irrelevant in the evaluation process of the whole formula. This idea was put forward by De Finetti [14], and the basic example of such an irrelevant state of truth is an *if-then* rule whose antecedent is false.
- **half-true**: This understanding of the third truth-value as to do with the modeling of natural language and the idea of fuzziness. Namely, not all concepts of natural language are Boolean, namely some concepts admit of borderline cases to which such concepts only apply partially. In other words, concepts have examples that are typical and other that are peripheral. For instance a sparrow is a typical bird while a penguin is an exceptional one. This view of concepts as inducing a 3-partition of the set of interpretations is due to Gentilhomme [25].

The extension of Boolean truth-tables to three-valued connectives should be driven by the meaning of the third truth-value, under the constraint of coinciding with Boolean tables on $\{T, F\}$. For instance, the *undefined* state will infect any proposition, a component of which has the corresponding truth-value (because the resulting function will take no value on such an undefined argument). In particular, for any Boolean connective f, $f(M, t) = f(t, M) = M$, i.e., M acts as an absorbing element of the truth-value algebra. It yields the weak truth-tables of Kleene.

In the case of the understanding of M as *unknown*, Kleene's rationale makes it clear that it stands for the set $\{T, F\}$ of truth-values, representing hesitancy between T and F. The natural way of extending Boolean truth-tables to account for this rationale is to apply the canonical extension $\tilde{f}$ of functions $f : U \to V$ to set-valued arguments, namely, $\tilde{f} : 2^U \to 2^V$ is such that $\tilde{f}(A) = \{f(u)|u \in A\} \subseteq V$. This process is an elementary form of uncertainty propagation. Applying this calculus to the 16 connectives of Boolean logic, in particular, negation, conjunction, disjunction and implication, yields Kleene strong truth tables.

If M is understood as *irrelevant*, it will serve as a semi-group identity for binary connectives, in particular for conjunction and disjunction. The conjunction and disjunction defined by the two properties that, (i) they coincide with Boolean truth-tables and that (ii) M is an identity, are unique. They are associative and differ only for arguments (T, F). The three-valued truth-table of quasi-conjunction is given in Table 1. Mathematically, operation & is known as a conjunctive idempotent *uninorm* [13]. It is a monotonically increasing semigroup operation on the ordered set $\{T > M > F\}$ with identity M, that coincides with conjunction on $\{T, F\}$. It is the most elementary example of such an operation. This quasi-conjunction is actually the one of Sobocinski's three-valued logic [39], that pioneered relevance logic [2].

&	F	M	T
F	F	F	F
M	F	M	T
T	F	T	T

Table 1. Conjunctive uninorm

In the fourth view of the third truth-value, there seems to be much more freedom in the choice of three-valued connectives, but for the requirement of their coinciding with Boolean connectives on $\{T, F\}$. Consider for instance the conjunction of two concepts like *cheap* and *small* for a car, that clearly admit of peripheral elements. Consider a not fully *cheap* car that is not fully *small*, to what extent should it be considered as *cheap and small*? One may for instance admit that for a conjunction the degree of truth $v(cheap\, and\, small) \leq \min(v(cheap), v(small))$:[1]. It stills leaves it open whether $t(M, M)$ should be F or M. This choice depends on how demanding we are regarding the conjunction. If we decide that the corresponding car should be deleted as possessing two defects, then $M \otimes M = F$ and we get Lukasiewicz conjunction (that he does not seem to have used explicitly, even though it can be derived from his implication). Being tolerant leads to using min, as a conjunction. These considerations lead to using Lukasiewicz 3-valued logic as a formal setting for the most elementary kind of gradual concepts. Indeed, both conjunctions can be retrieved from the Lukasiewicz implication.

[1] Although even this assumption can be questionable in practice

$\odot$	F	M	T
F	F	F	F
M	F	F	M
T	F	M	T

Table 2. Lukasiewicz conjunction

Note that in all the above understandings of the third truth-value M, the intuitively satisfactory negation connective is an involutive function ν on $\{T, M, F\}$, such that $\nu(M) = M$. Moreover, conjunctions and disjunctions will be related via ν by a De Morgan equality. This setting is clearly due to our emphasis on connectives, as opposed to using the notion of semantic deduction as a starting point (for instance, extending $\phi \models \psi$ as $v(\phi) \leq v(\phi)$ in terms of valuations).

3 Model-theoretic views of three-valued valuations

Many-valued logics traditionally use a standard propositional language that does not much differ from the one of Boolean logic. The difference lies in the choice of axioms and inference rules. Assume a propositional language $\mathcal{L}$ is used, with a set A of atomic propositions, with connectives of negation and conjunction, and let Ω be the set of classical interpretations $\omega : A \to \{T, F\}$. Let S be a set of states in which each proposition ϕ of the language can be evaluated on $\{T, M, F\}$. It leads us to attach to any proposition ϕ of the language a partition of S into three subsets $S_T(\phi), S_M(\phi), S_F(\phi)$: $S_t(\phi)$ contain those states in which ϕ has truth-value $t \in \{T, M, F\}$. The four views of the third truth-value M outlined above lead to different ways of interpreting states and of envisaging this three-partition of S.

- If M related to the idea of being undefined, one can choose $S = \Omega$ and one may consider for each atomic proposition a the subset $S_M(a)$ of interpretations, where a is undefined. Let $A(\phi)$ be the set of atoms appearing in ϕ. Then $S_M(\phi) = \cup_{a \in A(\phi)} S_M(a)$. Then any formula ϕ is a partially defined Boolean function: it is evaluated classically on each interpretation or state where it is defined. The set of models of ϕ is a subset of states in $S \setminus S_M(\phi)$.
- If M means unknown, the set $S_M(\phi)$ of states where the truth-value of ϕ is equal to *unknown* is then interpreted as the set of states where the (classical) truth-value of the classical proposition encoded by ϕ is unknown. Then the proposition ϕ has a set $[\phi]$ of classical models that is ill-known : $S_T(\phi) \subseteq [\phi] \subseteq S_T(\phi) \cup S_M(\phi)$. There may be several causes to such an incompleteness:
  1. The set S represents a set of states of incomplete knowledge. The actual state of the world is ill-known, and is only known to lie in a subset $E \subset \Omega$. In the line of Kleene's view, the epistemic set E is described by means of a partial valuation v consisting to declaring each atom $a \in A$ as being true, false or unknown. Then $E_v = \cap_{a:v(a)=T}[a] \setminus \cup_{a:v(a)=F}[a]$. In this setting, $S = \{E \subseteq \Omega | E = E_v$ for some $v\}$ and $S_T(\phi) = \{E_v | v \models_T \phi\}$, and $S_F(\phi) = \{E_v | v \models_F \phi\}$ where $\models_T$ and $\models_F$ are the satisfiability and falsification relation in partial logic [8]:
     - $v \models_T a$ if and only if a is true for v; $v \models_F a$ if and only if a is false for v;
     - $v \models_T \neg p$ if and only if $v \models_F p$; $v \models_F \neg p$ if and only if $v \models_T p$;
     - $v \models_T p \wedge q$ if and only if $v \models_T p$ and $v \models_T q$;

- $v \models_F p \wedge q$ if and only if $v \models_F p$ or $v \models_F q$
- $v \models_T p \vee q$ if and only if $v \models_T p$ or $v \models_T q$;
- $v \models_F p \vee q$ if and only if s $\models_F p$ and $v \models_F q$.

2. Another incomplete information scenario is that of rough sets. $S = \Omega$, but each formula ϕ refers to an ill-known subset of interpretations, with lower approximation $S_T(\phi)$ and upper approximation $S_T(\phi) \cup S_M(\phi)$. This situation is caused by the the existence of an equivalence R relation on Ω that prevents the accurate perception of interpretations, so that $S_T(\phi) = \cup_{R(\omega) \subseteq [\phi]} R(\omega)$, and $S_F(\phi) = \cup_{R(\omega) \subseteq [\neg\phi]} R(\omega)$, where $R(\omega)$ is the equivalence class containing ω. Banerjee and Chakraborty [4, 3] have tried to capture this situation using Lukasiewicz three-valued logic. The idea is to build a logic whose formulas ϕ are viewed as pairs $(\underline{S}(\phi), \overline{S}(\phi))$ of subsets of Ω under the following understanding of logical operations
   - $\vdash \phi \to \psi$ means $\underline{S}(\phi) \subseteq \underline{S}(\psi)$ and $\overline{S}(\phi) \subseteq \overline{S}(\psi)$ (rough implication);
   - $\phi \wedge \psi$ stands for the pair $(\underline{S}(\phi) \cap \underline{S}(\psi), \overline{S}(\phi) \cap \overline{S}(\psi))$;
   - $\phi \vee \psi$ stands for the pair $\underline{S}(\phi) \cup \underline{S}(\psi), \overline{S}(\phi) \cup \overline{S}(\psi))$;
   - $\neg\phi$ stands for the pair $(\overline{S}(\phi)^c, \underline{S}(\phi)^c)$.

   Banerjee [3] shows that the corresponding calculus is isomorphic to Łukasiewicz three-valued logic using his implication and negation connectives as primitives, letting $\phi \wedge \psi \equiv (\phi \to \psi) \to \psi$ and $\phi \vee \psi \equiv \neg(\neg\phi \wedge \neg\psi)$. Łukasiewicz's implication connective is such that $t(x \to y) = 1$ if $x \leq y$, $t(M \to F) = M$ and $t(T \to y) = y, \forall y \in \{T, M, F\}$, which accounts for rough implication. The lower approximation of ϕ is then syntactically expressed as $\neg(\phi \to \neg\phi)$.

– If the third truth value refers to the idea of *half-true* designating peripheral interpretations of a gradual proposition, then we can define $S = \Omega$, but, contrary to the previous case, the set of models of any formula ϕ *is* a fuzzy set with membership set $\{T, M, F\}$, that can be directly described as the pair $(S_T(\phi), S_M(\phi))$ made of central and peripheral interpretations. Łukasiewicz three-valued logic again looks like the most suitable setting for this view. Indeed, idempotent conjunctions $\wedge$ and disjunctions $\vee$ can be expressed as explained above, as well as non idempotent ones $\phi \oplus \psi = \neg\phi \to \psi$ and $\phi \otimes \psi = \neg(\phi \to \neg\psi)$.

4 Encapsulated logics

Three-valued logics mentioned previously share the use of a standard propositional language $\mathcal{L}$, but the semantic part of these logics appear not always so easy to grasp. While some views of M are purely a matter of defining what a proposition is (like in the fourth situation), other views involve epistemic aspects. While the syntactic part of the construction (the language) is very simple, the semantics is then quite complex. In this section we suggest that it is better to use a more expressive language where the epistemic and ontic aspects of the semantics are disentangled, what we call *encapsulated logics*. If there are several levels of semantics, there should be several levels of syntax, where degrees of truth in the higher-level language model some epistemic notion concerning the lower-level language. After pointing out some difficulties linked to a poor syntax, we briefly give a few examples of encapsulated logics.

For instance, in the case of the understanding of the third truth-value as *unknown*, there is a mix-up between epistemic and ontic notions: If we interpret M as $\{0, 1\}$ to justify the Kleene logic strong truth-tables, it means that we use Boolean truth-values T, F that can be called ontic and that define the essence of propositions that can be expressed in $\mathcal{L}$, and epistemic truth-values (as Belnap [7] calls them), $M = \{T, F\}$ and singletons $\{T\}, \{F\}$, that together form a Kleene algebra. It is well-known that Kleene logic has no tautologies. It has been argued elsewhere that this situation is anomalous [15]. If we assume a rational agent using a Boolean language (T, F as ontic truth-values), in the absence of specific knowledge, this agent should know the tautologies of the propositional calculus. Then, for instance, if ϕ and ψ are unknown, this agent should consider $\phi \wedge \neg\psi$ as unknown if ϕ and ψ are logically independent, but false if $\psi \equiv \neg\phi$. This conflicts with the use of truth-tables when computing the epistemic truth-value of a formulas in terms of the epistemic truth-values of its atoms. Worse, the use of truth tables enforces the use of partial valuations for representing incomplete knowledge (assigning epistemic truth-values to all atoms of the language). But only special kind of epistemic states, viewed as non-empty sets of mutually exclusive interpretations of the language, can be captured by partial valuations, i.e. those that can be described by consistent conjunctions of literals, not any disjunctions of interpretations. These two limitations imply that Kleene logic is a very approximate approach to the handling of epistemic uncertainty in propositional logic. The lack of tautologies indicates that in many cases a proposition that is arguably known as true or as false using constraint propagation techniques will be considered unknown via truth-tables.

In the case of the use of Lukasiewicz logics of rough sets, the difficulty posed by the approach has been studied in [12]. Namely, in order to properly justify the intersection of pairs of upper and lower approximations $(\underline{S}(\phi) \cap \underline{S}(\psi), \overline{S}(\phi) \cap \overline{S}(\psi))$ accounting for the conjunction connective, one must find a set of interpretations expressible in terms of the actual $[\phi]$ and $[\psi]$ the upper and lower approximations of which form the pair $(\underline{S}(\phi) \cap \underline{S}(\psi), \overline{S}(\phi) \cap \overline{S}(\psi))$. Banerjee and Chakraborty found such a set, of the form $([\phi] \cap [\psi]) \cup ([\phi] \cap \overline{S}(\psi) \cap \overline{S}(\chi)^c)$, where $[\chi] = [\phi] \cap [\psi]$. It is clear that the latter set depends on the equivalence relation R, which is the epistemic ingredient of the construction. Hence, the epistemic and ontic components are again entangled in this approach.

One way out of this kind of difficulty is to resort to encapsulated logics. The idea is as follows. Consider a language $\mathcal{L}_1$ as above. This is the ontic language. Define another language $\mathcal{L}_2$ whose atoms are built from formulas of $\mathcal{L}_1$. Namely, $\mathcal{L}_2$ is a first order language that uses one or several predicates P_i, referring to some epistemic notion of concern, and $\mathcal{L}_1$ is the set of constants that will be used for the arguments of these predicates. For instance, if the predicate is of arity 2, and $\phi, \psi \in \mathcal{L}_1$, then $P_i(\phi, \psi)$ is an atom of $\mathcal{L}_1$. $\mathcal{L}_2$ does not include formulas of $\mathcal{L}_1$. This why we can say that $\mathcal{L}_1$ is encapsulated into $\mathcal{L}_2$. On top of axioms of first order logic one, must add the encapsulation axiom ensuring that at the higher level, the lower language is properly used:

$$\forall j, \text{ if } \phi_i \equiv \psi_i, i = 1, \ldots, n \text{ then } P_j(\phi_1, \ldots, \phi_n) \equiv P_j(\psi_1, \ldots, \psi_n).$$

At the semantic level, this trick enables degrees of truth pertaining to ϕ to be distinct from degrees of truth of formulas of $\mathcal{L}_2$, for instance of $P(\phi)$ that may capture epistemic issues dealing with ϕ. This idea goes back to attempts by Hájek et al. [29] who capture probability theory inside a many-valued logic. In this case, $\mathcal{L}_1$ is propositional logic, while $\mathcal{L}_2$ contains only one unary predicate P (for *probable*) that is many-valued, and the degree of truth of $P(\phi)$ is interpreted as the degree of probability of the event encoded by the Boolean formula ϕ.

In the following, we give two examples of encapsulated logics disentangling ontic and epistemic truth-values of three-valued logic, one dealing with the truth-value *unknown*, one dealing with the truth-value *irrelevant*. In both cases, $\mathcal{L}_1$ is propositional logic.

4.1 The logic of *unknown*

The language $\mathcal{L}_2$ uses a unary predicate $\mathbf{T}$ and $\mathbf{T}(\phi)$ means that ϕ is known to be true (or believed: we do not deal here with the distinction between belief and knowledge). A formula $\Phi \in \mathcal{L}_2$ is either of the form $\mathbf{T}(\phi)$, or $\neg\Phi$ or $\Phi \wedge \Psi$.The following abbreviations are used:

– $\mathbf{F}(\phi)$ stands for $\mathbf{T}(\neg\phi)$ and means that ϕ is known to be false;
– $\mathbf{U}(\phi)$ is short for $\neg\mathbf{T}(\phi) \wedge \neg\mathbf{F}(\phi)$ and means that ϕ is unknown.

We thus clearly distinguish between the claim that ϕ is true from the claim that ϕ is known to be true. Moreover, the epistemic valuation *unknown* is explicitly modelled as ignorance about truth or falsity of ϕ by means of a derived predicate. Axioms should be added to those of first-order calculus (where $\rightarrow$ denotes material implication):

$(A1)$: $\mathbf{T}(\phi) \rightarrow \mathbf{T}(\psi)$, whenever $\vdash p \rightarrow q$.
$(A2)$: $\mathbf{T}(\phi \wedge \psi) \equiv \mathbf{T}(\phi) \wedge \mathbf{T}(\psi)$.
$(A3)$: $\mathbf{T}(\top)$.
$(A4)$: $\mathbf{T}(\phi) \rightarrow \neg\mathbf{F}(\phi)$.

(A1) enforces coherence of the epistemic valuation $\mathbf{T}$ with respect to the logical inference of proposition in the lower language. Axiom (A2) means that, for the source, asserting the truth of a conjunction of propositions in $\mathbf{L}$ is the same as asserting the truth of the conjuncts. Axiom (A3) assumes the source considers all classical tautologies as true. Axiom (A4) forbids the source from claiming simultaneously the truth of ϕ and its falsity. It means the knowledge is consistent.

It is not hard to be convinced that this logic is the same as a fragment of the KD system of modal logic, if we use $\Box p$ instead of $\mathbf{T}(\phi)$, as studied by Banerjee and Dubois [5].

Now, satisfiability in $\mathcal{L}_2$ is defined in a standard way, but the set of valuations of $\mathcal{L}_2$ is isomorphic to $\{E \subseteq \Omega, E \neq \emptyset\}$ (recall that due to axiom (A4), the knowledge is consistent). Indeed, suppose a valuation $v_2 : \{\mathbf{T}(\phi), \phi \in \mathcal{L}_1\} \rightarrow \{T, F\}$ Due to axioms A1-A4, the set $K(v_2) = \{\phi, v_2(\mathbf{T}(\phi)) = T\}$ is clearly a consistent deductively closed set of $\mathcal{L}_1$ formulas. Then there exists a non-empty subset $E(v_2) = [\phi] \in \Omega$ of interpretations of $\mathcal{L}_1$ such that $K(v_2) = \{\psi | \phi \models \psi$. Then the semantics of $\mathcal{L}_2$ can be expressed as follows: $v_2(\mathbf{T}(\phi)) = T$ if and only if $\forall\omega \in E(v_2), \omega(\phi) = T$.

The reader can check the soundness and completeness of this logic with respect to this semantics. All axioms are obviously in agreement with it.

Compared to Kleene logic, one may notice that :

1. the semantics involves all possible incomplete information states, not just those that can be captured by partial models.
2. The encapsulated approach preserves all classical tautologies. In particular, the law of excluded middle and the law of non-contradiction in the lower level language can be expressed in the higher-level language: $\vdash \mathbf{T}(\phi \vee \neg\phi)$ and $\vdash \mathbf{T}(\phi \wedge \neg\phi)$ hold.
3. The excluded fourth law holds in $\mathcal{L}_2$ that is, $\vdash \mathbf{T}(\phi) \vee \mathbf{U}(\phi) \vee \mathbf{F}(\phi)$ and its components are mutually exclusive.
4. If ϕ, ψ are logically independent propositions, $\mathbf{F}, \mathbf{F}' \in \{\mathbf{T}, \mathbf{U}, \mathbf{F}\}$ and function $f_\wedge$ interprets conjunction according to Kleene logic truth-table, then $\{\mathbf{F}(\phi), \mathbf{F}'(\psi)\} \vdash f_\wedge(\mathbf{F}, \mathbf{F}')(\phi \wedge \psi)$. except if $\mathbf{F} = \mathbf{F}' = \mathbf{U}$. So Kleene truth-tables are recovered when possible under the constraints induced by Boolean propositional logic $\mathcal{L}_1$.

While the interpretation of the truth-tables in terms of set-valued extensions of Boolean connectives indicates that Kleene logic is sound and will not produce wrong results (false propositions, where classical logic would predict tautologies), it is clearly insufficient to propagate incomplete knowledge and reason about ignorance in an optimal way.

Other encapsulated logics akin to the above are possibilistic logic [17] (where $\mathcal{L}_2$ is Gödel many-valued logic and only conjunction is used), and its multi-agent extensions [20], but also belief function logic [26].

4.2 The logic of irrelevance

Another example of encapsulated logic captures the case when a proposition can be irrelevant and as a consequence will not affect the truth-value of other propositions of which it is a component. This is the logic of exception-tolerant rules. The language $\mathcal{L}_2$ is then defined by means of a unique binary predicate $Imp(x, y)$. Terms x and y can be assigned formulas in $\mathcal{L}_1$, and $Imp(\phi, \psi)$ expresses the idea that in general ϕ implies ψ, up to exceptions. $Imp(\phi, \psi)$ encodes a rule *if* ϕ *then* ψ, and is a conditional. Only one connective is used : the conjunction denoted by &.

At the semantic level $\mathcal{L}_2$ is a three-valued logic and conjunction obeys the truth-table 1 of Sobocinski's logic [39]. Indeed, a rule *if* ϕ *then* ψ shares the set of interpretations Ω in 3 parts:

- *Examples of the rule*: interpretations where $\phi \wedge \psi$ is true.
- *Counterexamples of the rule*: interpretations where $\phi \wedge \neg\psi$ is true;
- *Irrelevant cases*: interpretations where ϕ is false.

This approach was originally proposed by De Finetti [14] who tried to model conditional events involved in conditional probabilities. It has been rediscovered several times in the literature (see bibliographies in the references [27, 19]), especially by [1], and [9]. Hence we define interpretations of $\mathcal{L}_2$ as:

- $v_2(Imp(\phi, \psi)) = T$ if and only if $\omega(\phi \wedge \psi) = T$,
- $v_2(Imp(\phi, \psi)) = F$ if and only if $\omega(\phi \wedge \neg\psi) = T$
- $v_2(Imp(\phi, \psi)) = I$ otherwise (we denote the third truth-value by I for *irrelevant.*

A conditional $Imp(\phi, \psi)$ can be viewed as a *set* of propositions in $\mathcal{L}_1$ rather than a single one, namely $\{\chi : \phi \wedge \psi \subseteq \chi \subseteq \neg\phi \vee \psi\}$. It forms an interval in the Boolean algebra of propositions, lower bounded by the conjunction $B \wedge A$ and upper bounded by material implication $\neg A \vee B$. This is the set of Boolean solutions χ to the equation $\psi \wedge \phi = \chi \wedge \phi$, i.e., a Boolean form of Bayes rule holds (just write $Imp(\phi, \psi)$ as "$\psi|\phi$" in the style of probability theory).

A rule $Imp(\phi, \psi)$ semantically implies another rule $Imp(\zeta, \chi)$, if the latter has more examples and less exceptions than the former:

$$Imp(\phi, \psi) \models Imp(\zeta, \chi) \text{ if and only if } \phi \wedge \psi \models \zeta \wedge \chi \text{ and } \zeta \wedge \neg\chi \models \phi \wedge \neg\psi.$$

This intuition is in agreement with the inequality of truth-values on $\mathcal{L}_2$: equipping the truth-set $\{T, F, I\}$ with the intuitive logical ordering $T > I > F$, the entailment between rules also reads:

$$Imp(\phi, \psi) \models Imp(\zeta, \chi) \text{ if and only if } v_2(Imp(\phi, \psi)) \leq v_2(Imp(\zeta, \chi)).$$

Noticing that $\zeta \wedge \neg\chi \models \phi \wedge \neg\psi$ equivalently writes $\phi \to \psi \models \zeta \to \chi$, this is also the same as rough entailment discussed above between pairs of nested propositions $(\phi \wedge \psi, \phi \to \psi)$ and $(\zeta \wedge \chi, \zeta \to \chi)$.

Now, consider a set Δ made of two rules $Imp(\phi_i, \psi_i), i = 1, 2$. We can extend ideas of example and counterexample to more then one rule.

- Δ is said to be verified by an $\mathcal{L}_1$-interpretation ω if ω is an example of at least one rule and is no counterexample of the other; this is naturally an example of Δ.
- The rule base Δ is falsified by an interpretation ω if ω is a counterexample of at least one rule in it; this is naturally a counterexample of Δ.
- An interpretation ω is irrelevant for the rule base Δ if it is irrelevant to both rules.

These requirements completely determine the conjunction $Imp(\phi_1, \psi_1)\&Imp(\phi_2, \psi_2)$ of the two rules:

- $v_2(Imp(\phi_1, \psi_1)\&Imp(\phi_2, \psi_2)) = F$,
  if $v_2(Imp(\phi_1, \psi_1)) = F$ or $v_2(Imp(\phi_2, \psi_2)) = F$;
- $v_2(Imp(\phi_1, \psi_1)\&Imp(\phi_2, \psi_2)) = I$,
  if $v_2(Imp(\phi_1, \psi_1)) = I$ and $v_2(Imp(\phi_2, \psi_2)) = I$;
- $v_2(Imp(\phi_1, \psi_1)\&Imp(\phi_2, \psi_2)) = T$ otherwise.

This is the uninorm (hence associative) of Table 1 precisely coinciding with the truth-table for conjunction of Sobocinski's logic and the following semantic equivalence can be checked :
$Imp(\phi_1, \psi_1)\&Imp(\phi_2, \psi_2) \equiv Imp(\phi_1 \wedge \phi_2, (\phi_1 \to \psi_1) \wedge (\phi_2 \to \psi_2))$
(where $\to$ is material implication). This is also the definition of quasi-conjunction QC in Adams [1] probabilistic conditional logic.

At this point we can define the semantic inference of rules from rules in $\mathcal{L}_2$ based on this three-valued semantics: A consistent set of rules Δ is said to semantically imply another rule $Imp(\phi, \psi)$, denoted by $\Delta \models Imp(\phi, \psi)$ if and only if

$$\exists \Sigma \subseteq \Delta, QC(\Sigma) \models Imp(\phi, \psi).$$

Note that this is not equivalent to $QC(\Delta) \models Imp(\phi, \psi)$ due to interference between rules.

It is possible to show that this three-valued semantic inference validates all postulates of system **P** of the non-monotonic logic of conditional assertions of Kraus et al. [33]. In fact, rewriting system P in $\mathcal{L}_2$ leads to the following additional axiom and inference rules:

- *Left Logical Equivalence*: if $\phi \equiv \psi$ then $\vdash Imp(\phi, \chi)$ if and only if $\vdash Imp(\psi, \chi)$ (LLE)
- *Right Weakening*: if $\psi \models \chi$ then $Imp(\phi, \psi) \vdash Imp(\phi, \chi)$ (RW)
- *Reflexivity Axiom*: $\vdash Imp(\phi, \phi)$
- *Left OR*: from $Imp(\phi, \chi)$ and $Imp(\psi, \chi)$ deduce $Imp(\phi \vee \psi, \chi)$ (LOR)
- *Cumulativity*: if $\vdash Imp(\phi, \psi)$ then $\vdash Imp(\phi, \chi)$ if and only $Imp(\phi \wedge \psi, \chi)$

However, it is easy to see that some usual deduction patterns fail in the three-valued logic $\mathcal{L}_2$:

- *Monotony*: $Imp(\phi, \psi) \not\models Imp(\phi \wedge \chi, \psi)$: indeed the latter rule has less examples than the first one.

– *Transitivity*: $\{Imp(\phi, \psi), Imp(\psi, \chi)\} \not\models Imp(\phi, \chi)$; indeed, an example to $Imp(\psi, \chi)$ that falsifies ϕ verifies $\{Imp(\phi, \psi), Imp(\psi, \chi)\}$ but it is not an example of the conclusion.
– *half deduction theorem*: $Imp(\phi, \neg\psi \vee \chi) \not\models Imp(\phi \wedge \psi, \chi)$; indeed models of $\phi \wedge \neg\psi$ verify the premise, not the conclusion.

It was proved [19] that system **P**, handling conditional statements of the form $Imp(\phi, \psi)$ (understood as reifications of non monotonic inference rules) is a syntactic counterpart of the three-valued logic of conditional events. From available results in the literature it follows that $\Delta \vdash_P Imp(\phi, \psi)$ if and only if $\Delta \models Imp(\phi, \psi)$. This logic is also closely related to Adams [1] probabilistic conditional logic, based on infinitesimal probabilities. More details on these connections can be found in Dubois and Prade [19, 21]. This logic actually provide symbolic underpinnings for reasoning with conditional probabilities as well, since if P is a positive probability distribution, then $Imp(\phi, \psi) \models Imp(\zeta, \chi)$ implies $P(\psi|\phi) \leq P(\zeta|\chi)$.

Other formalisms that use the notion of rule are logic programming formalisms that are often connected with three-valued logics. Casting them in the framework of encapsulated logics may shed light on their significance and their meaning (see [22]).

5 Conclusion

This short paper suggests that three-valued logics should be studied from the point of view of their pragmatics, not just from a syntactic and formal semantic point of view. Indeed, one may argue that in some sense three-valued logics are almost as basic as Boolean logic. A number of existing three-valued logic exists, that have been extensively studied in the past century, but there has been some confusion about what they can actually be useful for. Some bridges have been laid bare to uncertainty handling, rough sets, logic programming, reasoning with exceptions etc. However, as discussed here, these attempts, as interesting as they may be, are not perfect. One reason is that the mere use of a single propositional language with truth-functional semantics may be unable to capture all subtleties of a pragmatic notion, and to go beyond ontic truth-values. Here we have suggested the notion of encapsulated logics as a tool to separate epistemic notions from ontic ones at the semantic level. We show that the handling of ignorance with Kleene logic can be enhanced using an encapsulated approach akin to the modal logic. We also take the example of exception tolerant-rules where the three-valued calculus of Sobocinski looks instrumental.

There is a need to explore three-valued logics in a more systematic way using the approach outlined here, so as to establish each of them as a basic tool for exploring suitable pragmatic notions. The same methodology looks reasonable for other multiple -valued that aim at expressing more that ontic truth notions, like Belnap's four-valued logic for inconsistency handling [6, 7], and paraconsistent logics as well.

References

1. E. W. Adams. *The Logic of Conditionals*. D. Reidel, Dordrecht, The Netherlands, 1975.
2. A. Anderson and N. Belnap. *Entailment - The Logic of Necessity and Relevance*, volume 1. Princeton University Press, Princeton, N. J., 1975.

3. M. Banerjee. Rough sets and 3-valued lukasiewicz logic. *Fundam. Inform.*, 31(3/4):213–220, 1997.
4. M. Banerjee and M. K. Chakraborty. Rough sets through algebraic logic. *Fundam. Inform.*, 28(3-4):211–221, 1996.
5. M. Banerjee and D. Dubois. A simple modal logic for reasoning about revealed beliefs. In C. Sossai and G. Chemello, editors, *Proc. ECSQARU 2009, Verona, Italy, LNAI 5590*, pages 805–816. Springer-Verlag, 2009.
6. N. D. Belnap. How a computer should think. In G. Ryle, editor, *Contemporary Aspects of Philosophy*, page 3056. Oriel Press, 1977.
7. N. D. Belnap. A useful four-valued logic. In J.M. Dunn and G. Epstein, editors, *Modern Uses of Multiple-Valued Logic*, pages 8–37. D.Reidel, Dordrecht, The Netherlands, 1977.
8. S. Blamey. Partial logic. In D. Gabbay and F. Guentner, editors, *Handbook of Philosophical Logic, 2d Ed.*, volume 5, pages 261–353. Kluwer Academic Publ., Dordrecht, The Netherlands, 1998.
9. P. Calabrese. An algebraic synthesis of the foundations of logic and probability. *Information Sciences*, 42(3):187–237, 1987.
10. M. K. Chakraborty and Sanjukta Basu. Graded consequence and some metalogical notions generalized. *Fundam. Inform.*, 32(3-4):299–311, 1997.
11. M. K. Chakraborty and Soma Dutta. Graded consequence revisited. *Fuzzy Sets and Systems*, 161(14):1885–1905, 2010.
12. D. Ciucci and D. Dubois. Truth-functionality, rough sets and three-valued logics. In *ISMVL*, pages 98–103. IEEE Computer Society, 2010.
13. B. De Baets, J. C. Fodor, D. Ruiz-Aguilera, and J. Torrens. Idempotent uninorms on finite ordinal scales. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 17(1):1–14, 2009.
14. B. De Finetti. La logique de la probabilité. In *Actes du Congrès Inter. de Philosophie Scientifique (1935)*, pages IV1–IV9, Paris, France, 1936. Hermann et Cie Editions.
15. D. Dubois. On ignorance and contradiction considered as truth-values. *Logic Journal of the IGPL*, 16(2):195–216, 2008.
16. D. Dubois. Have fuzzy sets anything to do with vagueness. In C. Fermueller et al., editor, *Logical Models of Reasoning with Vague Information*, page to appear. Springer, 2011.
17. D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In D. M. Gabbay, C. J. Hogger, J. A. Robinson, and D. Nute, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 439–513. Oxford University Press, 1994.
18. D. Dubois and H. Prade. Can we enforce full compositionality in uncertainty calculi ? In *Proc. of the 12th National Conf. on Artificial Intelligence (AAAI'94), Seattle, WA, 31/07/94-04/08/94*, pages 149–154, 1994.
19. D. Dubois and H. Prade. Conditional objects as nonmonotonic consequence relationships. *Special issue on Conditional Event Algebra, IEEE Trans. on Systems, Man and Cybernetics*, 24(12):1724–1740, 1994.
20. D. Dubois and H. Prade. Toward multiple-agent extensions of possibilistic logic. In *Proc. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2007)*, pages 187–192. IEEE, 2007.
21. D. Dubois and H. Prade. Non-monotonic reasoning and uncertainty theories. In Gerhard Brewka, Victor W. Marek, and Miroslaw Truszczynski, editors, *Nonmonotonic Reasoning: Essays Celebrating its 30th Anniversary*, Mathematical Logic and Foundations, pages 141–176. College Publications, London, 2011.

22. D. Dubois, H. Prade, and S. Schockaert. Rules and meta-rules in the framework of possibility theory and possibilistic logic. *Scientia Iranica*, to appear, 2011.
23. K. Fine. Vagueness, truth and logic. *Synthese*, 54:235–259, 1975.
24. B. Ganter, G. Stumme, and R. Wille, editors. *Formal Concept Analysis, Foundations and Applications*, volume 3626 of *Lecture Notes in Computer Science*. Springer, 2005.
25. Y. Gentilhomme. Les ensembles flous en linguistique. *Cahiers de Linguistique Théorique et Appliquée*, 5:47–63, 1968.
26. L. Godo, P. Hájek, and F. Esteva. A fuzzy modal logic for belief functions. *Fundam. Inform.*, 57(2-4):127–146, 2003.
27. I. R. Goodman, H. T. Nguyen, and E. A. Walker. *Conditional Inference and Logic for Intelligent Systems*. North-Holland, Amsterdam, 1991.
28. P. Hájek. *Metamathematics of Fuzzy Logic*. Kluwer, Dordrecht, 1998.
29. P. Hájek, L. Godo, and F. Esteva. Fuzzy logic and probability. In P. Besnard and S. Hanks, editors, *UAI*, pages 237–244. Morgan Kaufmann, 1995.
30. J. Y. Halpern, R. Fagin, Y. Moses, and M.Y. Vardi. *Reasoning About Knowledge*. MIT Press (Revised paperback edition), 2003.
31. J. Hintikka. *Knowledge and Belief*. Cornell University Press, 1962.
32. S. C. Kleene. *Introduction to metamathematics*. North–Holland Pub. Co., Amsterdam, 1952.
33. S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.
34. J. Kyburg and H.E. Smokler, editors. *Studies in Subjective Probability*, Huntington, N.Y., 1980. Krieger Pub. Co.
35. J. Łukasiewicz. Philosophical remarks on many-valued systems of propositional logic (1930). In L. Borkowski, editor, *Selected Works*, pages 481–495. North-Holland, 1970.
36. Z. Pawlak. *Rough sets – Theoretical aspects of reasoning about data*. Kluwer Academic Publishers, Dordrecht, 1991.
37. Jayanta Sen and M. K. Chakraborty. A study of interconnections between rough and 3-valued lukasiewicz logics. *Fundam. Inform.*, 51(3):311–324, 2002.
38. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, N.J., 1976.
39. B. Sobociński. Axiomatization of a partial system of three-value calculus of propositions. *Journal of Computing Systems*, 1:23–55, 1952.
40. T. Williamson. *Vagueness*. Routledge, London, 1994.
41. L.A. Zadeh. Fuzzy sets as a basis for theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.

Didier Dubois
Institut de Recherche en Informatique deToulouse,
CNRS and Université de Toulouse, France.
`dubois@irit.fr`

A fragment of Mathematics from graded context: a proposal

Soma Dutta

1 Introduction

"You will never reach point B from point A as you must always get half-way there and half of the half, and half of that half and so on" – Zeno's Paradox

This is the intriguing nature of real numbers where we are dragged into. An infinite line holding all the real numbers along its length is lying on the ground. I thought from A I can reach to you at B. But each time I only could get *close* to you and *more close* to you and *more more close* to you. I even could not locate the point *next* to which you are standing.

While getting introduced to real numbers, we are taught – Real numbers are dense; between any two real numbers you can find infinitely many real numbers; howevermuch close you can imagine a real number to a particular real number x, there always exists a real number more close to x, etcetera. Almost all books of real numbers do use the phrases like, 'x tends to y', 'x is close to infinity (∞)', 'given any large positive number' etcetera, in their formal presentation for real numbers. Nevertheless, incorporating vagueness in its texture, standard presentation of real numbers makes us perceive it as a crisp theory.

Theoretically, real numbers are axiomatized as an ordered, algebraic structure $(R, \leq, +, \cdot)$ where $\leq$ is a crisp relation detrmining whether y lies right to x (or y is greater than x) or not, for any two real number x and y. So, according to the relation $\leq$, given any two numbers we can say which is right to which; but cannot locate an element lying just right to a particular number. This intrinsic continuous nature of real numbers opens a door to vagueness. In this article, one of the targets is to generate a language for real numbers where the basic predication is 'a number is close to another number'.

This enterprise of fuzzifying the language of reals prepares the ground for graded Mathematics.

There are already a number of approaches towards constructing an alternative Mathematics; one such found in 'Alternative Mathematics: the vague way' [1] by Jean Paul Van Bendegem. The paper talks about a Mathematics where a theorem like, 'small numbers has few prime factors' can be dealt with. So, numbers now have a fuzzy description and fuzzy quantification within its formal theory presentation. But graded Mathematics looks for something else. As in the existing approaches to fuzzy logic [5, 7, 8], in this approach too, to fuzzify Mathematics, researchers do not bother to address whether, if $S_1, S_2, \ldots, S_n, S$ are of matter of grade then in general 'S can be derived from $S_1, S_2, \ldots, S_n$' is also a matter of grade. So, from the context of graded consequence, after fuzzifying the object language for reals, an example of relevant question will be – to what extent one can derive 'y is large' from the premise consisting of 'x is large' and 'x is close to y'.

2 Graded consequence

We are now at the stage of revisiting the basic philosophy of the theory of graded consequence introduced by Chakraborty [2] in 1987. The theory was built on the idea that, if the object level formulas are many-valued, then it cannot be generally denied that the meta-level assertions are also many-valued. To elucidate the idea, let us observe the process of assigning grade to the notion of semantic consequence.

Let $\mathcal{L}$ be a language and $\{T_i\}_{i\in I}$ be a collection of fuzzy subsets [9] assigning grades/values from a lattice L to the formulas of $\mathcal{L}$. The notion of semantic consequence is represented by a meta-level sentence which states: 'α is a semantic consequence of X' if and only if 'for all T_i belonging to $\{T_i\}_{i\in I}$, if all members of X are true under T_i, then α is true under T_i'. $\ldots(\Sigma)$
That is, in the graded context, for all T_i, having the value of 'for all $\gamma \in X$, γ is true under T_i, i.e. $inf_{\gamma\in X} T_i(\gamma)$ and $T_i(\alpha)$, the value of 'α is true under T_i', we can compute the value of the 'if-then' statement, given in (Σ), by a many-valued meta-linguistic implication, say $\to_m$. And thus the grade to which α is semantic consequence of X i.e. $gr(X \mid\approx \alpha)$ becomes
$inf_i\{inf_{\gamma\in X} T_i(\gamma) \to_m T_i(\alpha)\}$ $\ldots(\Sigma')$.

This clarifies that the many-valuedness of object level formulas gets carried on to the meta-level concepts too. So, $\mid\approx$ is a L-fuzzy relation [4] representing the graded counterpart of the notion of semantic consequence relation. The operators 'inf', '$\to_m$' used in (Σ') are the respective 'infimum' and 'residuum of the multiplicative conjunction' of a complete residuated lattice L and they compute the meta-level 'for all' and 'if-then' respectively.

Initially, the idea of the theory of graded consequence was started by introducing a notion viz., $\mid\sim$, a graded consequence relation as a generalization of the classical consequence relation in many-valued context. Given a set of formulas X and a single formula α, $gr(X \mid\sim \alpha)$ determines the grade to which α is a consequence of X and is axiomatized by the following axioms.

(GC1) If $\alpha \in X$ then $\mathrm{gr}(X \mid\sim \alpha) = 1$ (Reflexivity)
(GC2) If $X \subseteq Y$ then $\mathrm{gr}(X \mid\sim \alpha) \leq \mathrm{gr}(Y \mid\sim \alpha)$ (Monotonicity)
(GC3) $inf_{\beta\in Y}\ \mathrm{gr}(X \mid\sim \beta) *_m \mathrm{gr}(X \cup Y \mid\sim \alpha) \leq \mathrm{gr}(X \mid\sim \alpha)$ (Cut)

Then the bridging between $\mid\sim$ and $\mid\approx$ was established [3] by the following theorems.

Theorem 2.1. *(Representation Theorem)*
(i) Given a collection of fuzzy subsets $\{T_i\}_{i\in I}$ *and a complete residuated lattice*

$(L, \to_m, *_m)$ for meta level algebra, $|\!\approx_{\{T_i\}_{i\in I}}$ is a graded consequence relation.
(ii) Given any graded consequence relation $|\!\sim$ there exists $\{T_i\}_{i\in I}$, a collection of fuzzy subsets over formulas, such that the fuzzy relation $|\!\approx_{\{T_i\}_{i\in I}}$ coincides with $|\!\sim$.

3 A fragment of Mathematics from graded context

In this section a language involving vague predicates will be introduced in such a way that one of its interpretations turns out to be the model of real numbers. Let us call the language as $\mathcal{L}_{\mathcal{R}}$.

Syntax of $\mathcal{L}_{\mathcal{R}}$
The language of $\mathcal{L}_{\mathcal{R}}$ is given as follows.

– Propositional constants: $\overline{0}$, $\overline{x}$, $\overline{y}$, $\overline{z}$, $\ldots -\overline{x}$, $-\overline{y}$, $-\overline{z}$, $\ldots$ (first kind)
$\overline{X}$, $\overline{Y}$, $\overline{Z}$, $\ldots$ (second kind)
– Variables: u, v, w, w'.
– Predicate symbols: cl_x, cl_y, cl_z, $\ldots$
$(+ve)cl_x$, $(+ve)cl_y$, $(+ve)cl_z$, $\ldots$
$(-ve)cl_x$, $(-ve)cl_y$, $(-ve)cl_z$, $\ldots$
L, Sm, $\in$, $=$
– Function symbol: $\oplus$, $\odot$ (each of arity 2)
$Supp_{(+ve)cl}$, $Supp_{(-ve)cl}$ (each of arity 1)
– Connectives: $\&$, $\vee$, $\neg$, $\to$, $\to_n$
– Quantifiers: $\forall$, $\exists$, $\exists_{inf}$
– Term: · All propositional constants and variables are terms.
· For t_1, t_2, t, the terms other than the second kind of propositional constants, $t_1 \oplus t_2$, $t_1 \odot t_2$, $Supp_{(+ve)cl}(t)$, $Supp_{(-ve)cl}(t)$ are terms.

Following are some types of atomic well-formed formulas (wffs) involving the basic predicates of $\mathcal{L}_{\mathcal{R}}$.

– Atomic wffs: · $cl_x(t_1^1)$, $\ldots$, $(+ve)cl_x(t_1^1)$, $\ldots$, $(-ve)cl_x(t_1^1)$, $\ldots$.
· $L(t_1^1)$, $\ldots$, $Sm(t_1^1)$, $\ldots$, $t_1^1 \in t_1^2$, $\ldots$, $t_1^1 = t_2^1$, $t_1^2 = t_2^2$,
· $t_1^1 \in Supp_{(+ve)cl}(t_2^1)$, $\ldots$, $t_1^1 \in Supp_{(-ve)cl}(t_2^1)$, $\ldots$

where t_1^1, t_2^1 are terms corresponding to first kind of propositional constants, variables or terms involving them and t_1^2, t_2^2 are used for second kind of propositional constants.

An interpretation of $\mathcal{L}_{\mathcal{R}}$
Let T be a model of $\mathcal{L}_{\mathcal{R}}$ consisting of a domain of interpretation, an interpretation function $\mathcal{I}$ and a value set $[0,1]$. The domain of interpretation for the propositional constants $\overline{0}, \overline{x}, \overline{y}, \overline{z}, \ldots$ is the set of all real numbers $0, x, y, z, \ldots$ (non-negative real numbers) and that of $-\overline{x}, -\overline{y}, -\overline{z}, \ldots$ are $-x, -y, -z, \ldots$ (negative real numbers) respectively. The domain of interpretation of the propositional constants of the second kind is the set of all sets of real numbers. The variables v, w, w' ranges over the propositional constants of the first kind and u ranges over the propositional constants of the form $\overline{x}$, $\overline{y}$, $\overline{z}$, $\ldots$.

Interpretation of the fuzzy predicates of the form cl_x i.e. $\mathcal{I}(cl_x)$ is given by a fuzzy relation, say $c_x : R \mapsto [0,1]$ such that
(i) $c_x(x) = 1$.
(ii) $c_x(y) \leq c_x(z)$ if $\mid y - x \mid \geq \mid z - x \mid$.
That is, for the wff $cl_x(\overline{y})$, $\mathcal{I}(cl_x(\overline{y}))$ determines $c_x(y)$, the value of the wff '$\overline{y}$ is close to $\overline{x}$'.

$\mathcal{I}((+ve)cl_x)$ is given by a fuzzy relation $\mathcal{I}((+ve)cl_x)$: $R \mapsto [0,1]$ such that
$\mathcal{I}((+ve)cl_x)(y) > 0$ if $y - x > 0$
$= 0$ if $y - x < 0$
$= 1$ if $y - x = 0$
$\mathcal{I}((+ve)cl_x)(y)$ is actually the truth value of the sentence $(+ve)cl_x(\overline{y})$.

$\mathcal{I}((-ve)cl_x)$ is given by a fuzzy relation $\mathcal{I}((-ve)cl_x)$: $R \mapsto [0,1]$ such that
$\mathcal{I}((-ve)cl_x)(y) > 0$ if $y - x < 0$
$= 0$ if $y - x > 0$
$= 1$ if $y - x = 0$
$\mathcal{I}((-ve)cl_x)(y)$ represents the truth value of the sentence $(-ve)cl_x(\overline{y})$.

$\mathcal{I}(L)$ and $\mathcal{I}(Sm)$ are given by two fuzzy relations say $\mathcal{I}(L) : R \mapsto [0,1]$ and $\mathcal{I}(Sm) : R \mapsto [0,1]$ respectively.

$\mathcal{I}(\overline{x} \in \overline{X})$ gets the truth value 1 if $\mathcal{I}(\overline{x})$ i.e. x is a member of $\mathcal{I}(\overline{X})$ i.e. X.

$\mathcal{I}(\oplus)$ and $\mathcal{I}(\odot)$ are the respective 'addition' $(+)$ and 'multiplication' $(\cdot)$ operations of reals.

Interpretation of $Supp_{(+ve)cl}$ and $Supp_{(-ve)cl}$ are functions from R to $P(R)$, defined by –
$\mathcal{I}(Supp_{(+ve)cl}(x)) = \{y/\mathcal{I}((+ve)cl_x(y)) > 0\}$ and
$\mathcal{I}(Supp_{(-ve)cl}(x)) = \{y/\mathcal{I}((-ve)cl_x(y)) > 0\}$.

Interpretation of connectives:
– Interpretation of & and $\vee$ are given by the lattice meet and join operators.
– Interpretation of $\neg$ is any standard complementation operation say, $\neg_a$ satisfying $\neg_a(1) = 0$ and $\neg_a(0) = 1$.
– $\rightarrow$ is interpreted by an operator $\Rightarrow$, defined by $a \Rightarrow b = 1$ if $a \leq b$
$= 0$ otherwise.
– $\rightarrow_n$ is interpreted by any standard fuzzy implication operator $\Rightarrow_n$ satisfying $a \Rightarrow_n b = 1$ iff $a \leq b$ and $1 \Rightarrow_n a = a$.

Interpretation of quantifiers for any predicate symbol P is given by:
– $\mathcal{I}(\forall u(P(u))) = inf_{\mathcal{I}(u)}\mathcal{I}(P(u))$.
– $\mathcal{I}(\exists u(P(u))) = sup_{\mathcal{I}(u)}\mathcal{I}(P(u))$.

– $\mathcal{I}(\exists_{inf})$ is given by a $[0,1]$-fuzzy subset over $P(R)$.
Here, it should be mentioned that as in general $\mathcal{I}(P)$ is a fuzzy relation, $\mathcal{I}(P(u))$ actually represents the degree of belongingness of $\mathcal{I}(u)$ to $\mathcal{I}(P)$.

Obtaining standard concepts of reals in the proposed fuzzified language $\mathcal{L}_{\mathcal{R}}$

After introducing the language $\mathcal{L}_{\mathcal{R}}$, the first task is to represent some basic concepts of real numbers within this fuzzified language. In order to do this, first some compound sentences will be framed in the proposed language and then it will be shown that their interpretation coincide with the usual understanding of some notions viz., 'neighbourhood', 'interior point', 'limit point', 'open set', 'closed set', 'bounded set' unbounded set', 'upper bound', 'lower bound' etcetera of real numbers.

§ Let for each propositional constant of the type $\overline{x}$, $\overline{y}$, $\overline{z}$, ... the language contains function symbols of the form $N_{\overline{x}}$, $N_{\overline{y}}$, $N_{\overline{z}}$ .... So, one can obtain terms of the form $N_{\overline{x}}(t')$ and hence expressions of the form $t \in N_{\overline{x}}(t')$ should be considered as a wff of the language, where t, t' are propositional constants of the first kind or a variable. Let us now represent a formula of the form $\overline{z} \in N_{\overline{y}}(\overline{x})$ in this proposed language. $\overline{z} \in N_{\overline{y}}(\overline{x})$ stands for the formula viz.,
$\big[(cl_x(\overline{x-y}) \to cl_x(\overline{z})) \vee (cl_x(\overline{x+y}) \to cl_x(\overline{z}))\big] \& \neg(\overline{z} = \overline{x-y}) \& \neg(\overline{z} = \overline{x+y})$.

As for each pair of real numbers x, y both $x+y$ and $x-y$ are real numbers, the langugae must contain propositional constants, say $\overline{x+y}$, $\overline{x-y}$ representing $x+y$, $x-y$ respectively. Now given the model T of $\mathcal{L}_{\mathcal{R}}$ we will show that the wff $\overline{z} \in N_{\overline{y}}(\overline{x})$ either gets the truth value 1 or 0.
$T(\overline{z} \in N_{\overline{y}}(\overline{x})) = 1$ if $T(\neg(\overline{z} = \overline{x-y})) = 1$, $T(\neg(\overline{z} = \overline{x+y})) = 1$ and
$T(cl_x(\overline{x-y}) \to cl_x(\overline{z})) = 1$ or $T(cl_x(\overline{x+y}) \to cl_x(\overline{z})) = 1$.
Otherwise, as '=' represents the crisp equality relation and the interpretation of '$\to$' is given by a $\{0,1\}$-valued function $\Rightarrow$, the value of $\overline{z} \in N_{\overline{y}}(\overline{x})$ will be 0.

Now $T(\overline{z} \in N_{\overline{y}}(\overline{x})) = 1$ implies $T(\overline{z} = \overline{x+y}) = 0$, $T(\overline{z} = \overline{x-y}) = 0$, and $\mathcal{I}(cl_x(\overline{x-y})) \leq \mathcal{I}(cl_x(\overline{z}))$ or $\mathcal{I}(cl_x(\overline{x+y})) \leq \mathcal{I}(cl_x(\overline{z}))$. Here we should note that interpretation of a closed formula/sentence gives the truth value of the formula in the model T. So, by writing $T(\alpha)$ we actually mean $\mathcal{I}(\alpha)$ for any closed formula α. Now, from the first line of this paragraph we can see that $T(\overline{z} \in N_{\overline{y}}(\overline{x})) = 1$ implies $z \neq x+y$, $z \neq x-y$ and $\mid z-x \mid \leq \mid y \mid$. That is, $\mid z-x \mid < \mid y \mid$. This inequality can be, in other way, written as $- \mid y \mid < z-x < \mid y \mid$, where y is the positive real number interpreting the propositional constant $\overline{y}$. Hence we obtain the definition of 'neighbourhood (nbd)' viz., for any $y > 0$, z belongs to the y-nbd of x if $x-y < z < x+y$.

So, $\mathcal{I}(N_{\overline{y}}(\overline{x}))$ is a subset of R, given by $\mathcal{I}(N_{\overline{y}}(\overline{x})) = \{z \in R / x-y < z < x+y\}$.

§ Let the language contain a function symbol 'Int' generating terms of the form $Int(\overline{X})$ in the language. So, expression of the form $t \in Int(\overline{X})$ (where t represents propositional constants of the first kind or variable) is a wff of the language. Now, in the language the wff $\overline{x} \in Int(\overline{X})$ is represented by the follow-

ing formula.
$\overline{x} \in Int(\overline{X}) \equiv \overline{x} \in \overline{X}$ & $\exists u(\forall v(v \in N_u(\overline{x}) \rightarrow v \in \overline{X})$.

As u runs over the propositional constants corresponding to the positive real numbers $x, y, z, \ldots$, N_u represents any one of $N_{\overline{x}}$, $N_{\overline{y}}$, $N_{\overline{z}}$, ... Here again, the proposed interpretation of the predicate symbol $\in$, connectives &, $\rightarrow$ and that of $N_u(\overline{x})$ imply that in the model T, $T(\overline{x} \in Int(\overline{X}))$ either gets the value 1 or 0.

Now, $T(\overline{x} \in Int(\overline{X})) = 1$ implies $T(\overline{x} \in \overline{X}) = 1$ and
$T(\exists u(\forall v(v \in N_u(\overline{x}) \rightarrow v \in \overline{X})) = sup_{\mathcal{I}(u)}(inf_{\mathcal{I}(v)}(\mathcal{I}(v \in N_u(\overline{x})) \Rightarrow \mathcal{I}(v \in \overline{X})))$
is 1. That is, there is some $\mathcal{I}(u)$ such that for all $\mathcal{I}(v)$, $\mathcal{I}(v \in N_u(\overline{x})) \Rightarrow \mathcal{I}(v \in \overline{X})$ gets the value 1. That is, the condition for $T(\overline{x} \in Int(\overline{X})) = 1$ turns out to be 'x is a member of X and for some interpretation $\mathcal{I}(u)$ of u there exists a $\mathcal{I}(u)$-nbd of x (i.e. $\mathcal{I}(N_u(\overline{x}))$ such that for all interpretation of v, $\mathcal{I}(v)$ belongs to $\mathcal{I}(u)$-nbd of x implies $\mathcal{I}(v)$ belongs to X'. This coincides with the standard definition of 'x is an interior point of X'.

So, interpretation of $Int(\overline{X})$ becomes –
$\mathcal{I}(Int(\overline{X})) = \{x \in X/\mathcal{I}(u)\text{-nbd of } x \subseteq X, \text{ for some } \mathcal{I}(u)\}$
$\qquad = \{x \in X/N_u(x) \subseteq X, \text{ for some } \mathcal{I}(u)\}$.

For convenience of writing, let us call $\mathcal{I}(N_u(\overline{x}))$ i.e. $\mathcal{I}(u)$-nbd of x as $N_u(x)$ and $\mathcal{I}(Int(\overline{X}))$ as $\mathrm{Int}(X)$.

§ Now we introduce the sentence describing $Open(\overline{X})$, where '$Open$' is assumed to be an unary predicate giving formulas of the form $Open(\overline{X})$ in the language.
$Open(\overline{X}) \equiv \forall v(v \in \overline{X} \rightarrow v \in Int(\overline{X}))$
In this case also, the proposed interpretation of $\in$, $\rightarrow$ and $Int(\overline{X})$ makes it evident that in the model T, $T(Open(\overline{X}))$ is either 1 or 0. Now $T(Open(\overline{X})) = 1$ means $inf_{\mathcal{I}(v)}(\mathcal{I}(v \in \overline{X}) \Rightarrow \mathcal{I}(v \in Int(\overline{X}))) = 1$. That is, for all interpretation of v, $\mathcal{I}(v)$ is a member of X implies $\mathcal{I}(v)$ is a member of $\mathrm{Int}(X)$; or in other words, $X \subseteq \mathrm{Int}(X)$. So, holding of the formula '$Open(\overline{X})$' in $\mathcal{L}_{\mathcal{R}}$ algebraically gives our standard definition of 'X is open'.

§ Let us now introduce a binary predicate symbol 'Lim', generating formulas of the form $Lim(\overline{x}, \overline{X})$ in the language $\mathcal{L}_{\mathcal{R}}$. $Lim(\overline{x}, \overline{X})$ stands for a wff in $\mathcal{L}_{\mathcal{R}}$ and it is given as follows.
$Lim(\overline{x}, \overline{X}) \equiv \forall u(\exists v(v \in N_u(\overline{x}) \rightarrow v \in \overline{X})$ &$\neg(v = \overline{x}))$

Again from the proposed interpretation of $\in$, $\rightarrow$, $=$ and $N_u(\overline{x})$ we can conclude that in the model T, $Lim(\overline{x}, \overline{X})$ either gets the value 1 or 0. Now, $T(Lim(\overline{x}, \overline{X})) = 1$ implies
$inf_{\mathcal{I}(u)}(sup_{\mathcal{I}(v)}(\mathcal{I}(v \in N_u(\overline{x})) \Rightarrow \mathcal{I}(v \in \overline{X})) \wedge \mathcal{I}(\neg(v = \overline{x}))) = 1$.
That is, $Lim(\overline{x}, \overline{X})$ gets the value 1 means that for all interpretation $\mathcal{I}(u)$ of u, there is some interpretation of v distinct from x for which $\mathcal{I}(v)$ belongs to $\mathcal{I}(u)$-nbd of x implies $\mathcal{I}(v)$ belongs to X. In more concise form we can represent the above condition as: $Lim(\overline{x}, \overline{X})$ holds if for all interpretation of u,
$[N_u(x) - \{x\}] \cap X \neq \phi$. This reminds our usual definition of 'x is a limit point of X'.

§ Now with the help of the predicate 'Lim' we shall represent a formula viz., $Closed(\overline{X})$ in the language of $\mathcal{L}_{\mathcal{R}}$. Let $Closed$ be a predicate symbol giving rise to formulas of the form '$Closed(\overline{X})$'. The formula representing $Closed(\overline{X})$ is given below.

$Closed(\overline{X}) \equiv \forall v(Lim(v, \overline{X}) \rightarrow v \in \overline{X})$.

Algebraically, in the model T, $Closed(\overline{X})$ gets either the value 1 or 0. That is, $inf_{\mathcal{I}(v)}(\mathcal{I}(Lim(v, \overline{X}) \Rightarrow \mathcal{I}(v \in \overline{X})) = 1$. This indicates that $Closed(\overline{X})$ holds if for all interpretation of v, $\mathcal{I}(v)$ is a limit point of X, then $\mathcal{I}(v)$ is a member of X. This is nothing but our standard definition of 'X is closed'.

§ Let $\mathcal{L}_{\mathcal{R}}$ contain a predicate symbol $Ubdd_a$ which generates 'X is unbounded above' by the formula $Ubdd_a(\overline{X})$. In $\mathcal{L}_{\mathcal{R}}$, $Ubdd_a(\overline{X})$ is represented as follows.

$Ubdd_a(\overline{X}) \equiv \forall v(v \in \overline{X} \rightarrow \exists w(\neg(w = v) \;\&\; w \in \overline{X} \;\&\; w \in Supp_{(+ve)cl}(v)))$

From the proposed interpretation of $Supp_{(+ve)cl}$, it is clear that for any interpretation of v, $Supp_{(+ve)cl}(v)$ generates an ordinary set of R. So, for any interpretation of v and w, $\mathcal{I}(w \in Supp_{(+ve)cl}(v))$ always gets either 1 or 0. So, like earlier cases, here also we can say that in the model T, $Ubdd_a(\overline{X})$ either gets the value 1 or 0. Now $T(Ubdd_a(\overline{X})) = 1$ means

$inf_{\mathcal{I}(v)}(\mathcal{I}(v \in \overline{X}) \Rightarrow sup_{\mathcal{I}(w)}(\mathcal{I}(\neg(w = v)) \wedge \mathcal{I}(w \in \overline{X}) \wedge \mathcal{I}(w \in Supp_{(+ve)cl}(v))))$

gets the value 1. That is, $Ubdd_a(\overline{X})$ holds if for all interpretation of v, $\mathcal{I}(v)$ belongs to X implies that there is some $\mathcal{I}(w)$ distinct from $\mathcal{I}(v)$, for which $\mathcal{I}(w)$ belongs to X as well as $\mathcal{I}(w)$ belongs to $\mathcal{I}(Supp_{(+ve)cl}(v))$.

Now, the definition of $\mathcal{I}(Supp_{(+ve)cl})$ implies that for any $\mathcal{I}(v)$, $\mathcal{I}(Supp_{(+ve)cl}(v))$ actually determines all those real numbers which are lying right to the number referred to by $\mathcal{I}(v)$. That is, given any two $\overline{x}$, $\overline{y}$, if $\overline{y} \in Supp_{(+ve)cl}(\overline{x})$ holds, then $y \geq x$.

So, in other words, from semantic angle '$Ubdd_a(\overline{X})$ holds' can be stated as: for all interpretation of v, $\mathcal{I}(v)$ belongs to X implies that there is some $\mathcal{I}(w)$ distinct from $\mathcal{I}(v)$, for which $\mathcal{I}(w)$ belongs to X and $\mathcal{I}(w) \geq \mathcal{I}(v)$. This reflects our standard definition of 'X is unbounded above'.

§ Similarly, introducing a predicate '$Ubdd_b$' in the language, 'X is unbounded below' can be represented in $\mathcal{L}_{\mathcal{R}}$. The formula representing $Ubdd_b(\overline{X})$ is given as follows.

$Ubdd_b(\overline{X}) \equiv \forall v(v \in \overline{X} \rightarrow \exists w(\neg(w = v) \;\&\; w \in \overline{X} \;\&\; w \in Supp_{(-ve)cl}(v)))$

From the interpretation of the predicate $(-ve)cl$ and hence $Supp_{(-ve)cl}$, it is clear that, if for some interpretation of w and v, $\mathcal{I}(w \in Supp_{(-ve)cl}(v)) = 1$, then $\mathcal{I}(w)$ lies left to $\mathcal{I}(v)$ i.e. $\mathcal{I}(w) \leq \mathcal{I}(v)$. So, holding $Ubdd_b(\overline{X})$ algebraically represents that for all interpretation of v, $\mathcal{I}(v)$ belongs to X implies that there is some interpretation of w distinct from $\mathcal{I}(v)$ for which $\mathcal{I}(w)$ belongs to X and $\mathcal{I}(w) \leq \mathcal{I}(v)$. This coincides with the standard definition of 'X is unbounded below'.

§ Let us now consider Bdd to be a predicate generating formulas of the type $Bdd(\overline{X})$ in $\mathcal{L}_{\mathcal{R}}$. $Bdd(\overline{X})$ is represented in the language $\mathcal{L}_{\mathcal{R}}$ by the following for-

mula.
$\exists v \exists w (\forall w' (w' \in \overline{X} \rightarrow ((v \in Supp_{(+ve)cl}(w')) \;\&\; (w \in Supp_{(-ve)cl}(w')))))$

So, with the interpretation of $Supp_{(+ve)vl}$ and $Supp_{(-ve)vl}$, it is now clear that $Bdd(\overline{X})$ holds if there are some interpretation of v, w such that for all interpretation of w', $\mathcal{I}(w')$ belongs to X implies $\mathcal{I}(v) \geq \mathcal{I}(w')$ and $\mathcal{I}(w) \leq \mathcal{I}(w')$. This conforms with the usual definition of 'X is bounded'.

§ Let us now introduce the binary predicate $U - bd$ in $\mathcal{L}_{\mathcal{R}}$. We shall now represent the formula viz., $U - bd(\overline{x}, \overline{X})$ in $\mathcal{L}_{\mathcal{R}}$.
$U - bd(\overline{x}, \overline{X}) \equiv \forall v (v \in \overline{X} \rightarrow \overline{x} \in Supp_{(+ve)cl}(v))$

So, it is quite obvious from the above expression that $U - bd(\overline{x}, \overline{X})$ holds if for all interpretation of v, $\mathcal{I}(v)$ belongs to X implies $\mathcal{I}(v) \leq x$. Thus holding of $U - bd(\overline{x}, \overline{X})$ in $\mathcal{L}_{\mathcal{R}}$ gives us the standard condition for 'x is an upper bound of X'.

§ After introducing $U - bd(\overline{x}, \overline{X})$ in $\mathcal{L}_{\mathcal{R}}$, we now introduce $Lub(\overline{x}, \overline{X})$ in $\mathcal{L}_{\mathcal{R}}$.
$Lub(\overline{x}, \overline{X}) \equiv U - bd(\overline{x}, \overline{X}) \;\&\; \forall v (U - bd(v, \overline{X}) \rightarrow v \in Supp_{(+ve)cl}(\overline{x}))$

That is, $Lub(\overline{x}, \overline{X})$ holds if x is an upper bound of X and for all interpretation $\mathcal{I}(v)$ of v, if $\mathcal{I}(v)$ is an upper bound of X then $\mathcal{I}(v) \geq x$. This reminds us the definition of 'x is the least upper bound of X'.

§ Next, we present 'x is a lower bound of X' in $\mathcal{L}_{\mathcal{R}}$. In order to do this, we introduce a predicate $L - bd$ in the language $\mathcal{L}_{\mathcal{R}}$. Now the formula $L - bd(\overline{x}, \overline{X})$ is represented by the following formula.
$L - bd(\overline{x}, \overline{X}) \equiv \forall v (v \in \overline{X} \rightarrow \overline{x} \in Supp_{(-ve)cl}(v))$

That is, $L - bd(\overline{x}, \overline{X})$ holds if for all interpretation $\mathcal{I}(v)$ of v, $\mathcal{I}(v)$ belongs to X then $x \leq \mathcal{I}(v)$. Thus the standard meaning of 'x is a lower bound of X' is captured in the language of $\mathcal{L}_{\mathcal{R}}$.

§ Now, with the help of the formula $L - bd(\overline{x}, \overline{X})$ we can introduce the notion $Glb(\overline{x}, \overline{X})$ representing 'x is the greatest lower bound of X'.
$Glb(\overline{x}, \overline{X}) \equiv L - bd(\overline{x}, \overline{X}) \;\&\; \forall v (L - bd(v, \overline{X}) \rightarrow \overline{x} \in Supp_{(+ve)cl}(v))$

Hence, $Glb(\overline{x}, \overline{X})$ holds if x is a lower bound of X and for all interpretation $\mathcal{I}(v)$ of v, if $\mathcal{I}(v)$ is a lower bound of X, then $\mathcal{I}(v) \leq x$. Thus, the definition of 'x is the greatest lower bound of X' is represented in the language $\mathcal{L}_{\mathcal{R}}$.

All these notions represented above are the standard, often-used notions of real numbers. The difference lies in the presentation of $\mathcal{L}_{\mathcal{R}}$, where the basic ingredient has been the idea of fuzzifying the concept viz., 'x is close to y'. And then the entire enterprise has been built on the basic fuzzy predicates of the type 'cl_x'. Also, it has been shown, that the crisp relation $\leq$ which is assumed to be the basic relation in the classical presentation of real numbers, can be obtained as a derived notion in $\mathcal{L}_{\mathcal{R}}$. So, we are now at the stage of representing all the axioms of $(R, \leq, +, \cdot)$, our known complete, ordered field of real numbers.

1. Algebraic axioms of $\mathcal{L}_{\mathcal{R}}$

(i) $\overline{x} \oplus \overline{y} = \overline{x+y}$
(ii) $\overline{x} \oplus (\overline{y} \oplus \overline{z}) = (\overline{x} \oplus \overline{y}) \oplus \overline{z}$
(iii) $\overline{x} \oplus \overline{0} = \overline{x}$
(iv) $\overline{x} \oplus \overline{-x} = \overline{0}$
(v) $\overline{x} \oplus \overline{y} = \overline{y} \oplus \overline{x}$
(vi) $\overline{x} \odot \overline{y} = \overline{x \cdot y}$
(vii) $\overline{x} \odot (\overline{y} \odot \overline{z}) = (\overline{x} \odot \overline{y}) \odot \overline{z}$
(viii) $\overline{x} \odot \overline{1} = \overline{x}$
(ix) $\neg(\overline{x} = \overline{0}) \rightarrow (\overline{x} \odot \overline{x}^{-1} = \overline{1})$
(x) $\overline{x} \odot \overline{y} = \overline{y} \odot \overline{x}$
(xi) $\overline{x} \odot (\overline{y} \oplus \overline{z}) = (\overline{x} \odot \overline{y}) \oplus (\overline{x} \odot \overline{z})$

In the presentation of the language of $\mathcal{L}_{\mathcal{R}}$, for not making the presentation cumbersome, we have not explicitly used propositional constants of the type $\overline{x}^{-1}$. But, for the convenience of understanding, here we have used $\overline{x}^{-1}$ as the propositional constant representing the real number which is the inverse of x.

From the proposed interpretation of $\oplus$, $\odot$, $\rightarrow$ and propositional constants, it is not difficult to check that all these formulas are true i.e. gets the value 1 in the model T.

2. Order axioms of $\mathcal{L}_{\mathcal{R}}$

(i) $((\overline{x} \in Supp_{(+ve)cl}(\overline{y}))$ & $(\overline{y} \in Supp_{(+ve)cl}(\overline{z}))) \rightarrow (\overline{x} \in Supp_{(+ve)cl}(\overline{z}))$
(ii) $(\overline{x} \in Supp_{(+ve)cl}(\overline{y})) \vee (\overline{y} \in Supp_{(+ve)cl}(\overline{x}))$
(iii) $(\overline{x} \in Supp_{(+ve)cl}(\overline{y})) \rightarrow (\overline{x} \oplus \overline{z} \in Supp_{(+ve)cl}(\overline{y} \oplus \overline{z}))$
(iv) $((\overline{x} \in Supp_{(+ve)cl}(\overline{0}))$ & $(\overline{y} \in Supp_{(+ve)cl}(\overline{z}))) \rightarrow (\overline{x} \odot \overline{y} \in Supp_{(+ve)cl}(\overline{x} \odot \overline{z}))$

From the proposed interpretation of $Supp_{(+ve)cl}$, we know $\overline{x} \in Supp_{(+ve)cl}(\overline{y})$ algebraically represents $y \leq x$. Hence we can observe that (i) represents the transitivity condition, (ii) represents the law of trichotomy and, (iii) and (iv) represent the law of compatibility of the ordered set $(R, \leq)$. And that the axioms viz., 2 (i), (ii), (iii), (iv) are true in the model T is quite obvious from the interpretation of $\in$, $\rightarrow$, &, $\vee$ and $Supp_{(+ve)cl}$.

3. Completeness axiom of $\mathcal{L}_{\mathcal{R}}$

$(\neg(\overline{X} = \overline{\phi})$ & $Bdd(\overline{X})) \rightarrow \exists v Lub(v, \overline{X})$

The validity of the completeness axiom is also reflected from the proposed interpretation of the logical symbols present in the above mentioned formula.

The meta-logic of classical Mathematics follows the laws of first order predicate logic. The opeartors computing the connectives of $\mathcal{L}_{\mathcal{R}}$ operate on many-valued components. But restricted to the set $\{0, 1\}$, all these operators for the connectives present in $\mathcal{L}_{\mathcal{R}}$ turn out to be exactly those of classical operators for 'implication', 'conjunction', 'disjunction' and 'negation'. The concepts of real numbers presented above viz., 'z belongs to y-nbd of x', 'x belongs to interior of X', 'X is open', 'x is a limit point of X', 'X is closed', 'X is unbounded above', 'X is unbounded below', 'X is bounded', 'x is an upper bound of X', 'x is the least upper bound of X', 'x is a lower bound of X', 'x is the greatest lower bound of X' are introduced in $\mathcal{L}_{\mathcal{R}}$ as two-valued sentences. So, while dealing with these notions, the meta-logic in this new framework also remains the same as that of classical one. Now as in $\mathcal{L}_{\mathcal{R}}$ all the axioms of real numbers can be presented, it is expected, concerning the above mentioned notions, that if α can be derived from X in the standard presentation of the formal theory of real numbers, then α can be obtained from X in $\mathcal{L}_{\mathcal{R}}$ also.

4 Going beyond the classical theory of reals

Though the basic language of $\mathcal{L}_{\mathcal{R}}$ has been fuzzified, till now the basic notions of real numbers have been presented crisply, as it is prevalently used in the existing literature of real number theory. But we do feel that in the real numbers, there are some concepts whose inherent vagueness need to be addressed distinctly and kept intact.

Let us consider the standard definition of 'x is close to ∞'. Existing definition states that 'given any large positive number G, x is always found to be greater than G'. But do we understand the phrase 'given any large positive number' as we understand an 'yes-no' concept? Definitely, we do not. But, still, these existing approaches make us perceive 'x is close to ∞' crisply. We would like to perceive 'x is close to ∞' as a matter of grade and hence in $\mathcal{L}_{\mathcal{R}}$, assuming cl_∞ as a predicate symbol, the concept is presented as follows.
$cl_\infty(\overline{x}) \equiv \forall u(L(u) \rightarrow_n (+ve)cl_u(\overline{x}))$

In the proposed interpretation, L and $(+ve)cl$ have been presented as fuzzy relations. Also $\rightarrow_n$ is computed by a fuzzy implication operator viz., $\Rightarrow_n$ which as an output may give a non-zero, non-unit value to $cl_\infty(\overline{x})$. So, generally, in $\mathcal{L}_{\mathcal{R}}$, $cl_\infty(\overline{x})$ is a graded concept. In the classical presentations, $L(u)$ and $(+ve)cl_u(\overline{x})$ are interpreted in a crisp way. That is, one can partition the domain of u in two halves; one consists of those interpretations of u for which $\mathcal{I}(L(u))$ is 0 and the other consists of those interpretations of u for which $\mathcal{I}(L(u))$ is 1. On the other hand, for every interpretation of u, $(+ve)cl_u$ is crisply interpreted by the set of those real numbers which lie right to the number, referred to by u. Now 'x is close to ∞'/'x tends to ∞' means, x is assumed to be right of any number. The phrase 'large number' does not play any role actually. And hence for any $\mathcal{I}(u)$, $(+ve)cl_u(\overline{x})$ always gets the value 1. So, for such an x, whatever be the $\mathcal{I}(u)$, $inf_{\mathcal{I}(u)}(\mathcal{I}(L(u)) \Rightarrow_n \mathcal{I}((+ve)cl_u(\overline{x})))$ always gets the value 1.

That is, to keep the concept 'x is close to ∞' as a two-valued notion, one only needs to interpret 'L' and '$(+ve)cl$' crisply. Now, let us explore, what kind of Mathematics can be developed if one allows 'x is close to ∞' i.e. $cl_\infty(\overline{x})$ as a many-valued notion. In the context of graded consequence, given the language $\mathcal{L}_{\mathcal{R}}$ and its interpretation in the model T, one can now be curious to know to what grade from the premise $cl_\infty(\overline{x})$ and $cl_x(\overline{y})$, $cl_\infty(\overline{y})$ can be derived. That is, calculating $gr(\{cl_\infty(\overline{x}), cl_x(\overline{y})\} \mid\approx cl_\infty(\overline{y}))$ and/or $gr(\{cl_\infty(\overline{x}), cl_x(\overline{y})\} \mid\approx L(\overline{y}))$ and/or $gr(\{cl_\infty(\overline{x}), L(\overline{y})\} \mid\approx cl_x(\overline{y}))$ and/or finding their interrelations may add a new dimension to the study of Mathematics.

There is one more thread in classical Mathematics, which may lead to vagueness. The thread is, how do we perceive 'infinity'. As Aristotle remarked [6],

"...generally the infinite is as follows: there is always another and another to be taken. And the thing taken will always be finite, but always different."

That is, 'infinity' means that *there is always something outside of what has been taken* [6]. We do perceive 'infinity' something as 'very large' - larger than anything one can imagine or percieve; nothing else. In [6], Jonathan Lear, described Aristotle's position about infinite number by the following line.
"...on his account, the most natural conception of the infinite by addition - that it is possible to exceed any given length - is impossible. Thus one cannot assign a numerical unit (I) to a standard unit length and hope that an infinity of numbers will be guaranteed by the fact that this length can be repeatedly added without end."

According to [6], to the intuitionist, the claim that a man may actually start by making theoretical division and given any number of divisions he has made, he could make another, gives nothing but an idea of *vaguely determined totality*. Lear thinks that it is absurd to suppose that some particular division is the last division one could make, and hence the number of the theoretical division one could actually carry out specifies only a vague totality.

How do we perceive 10^{100}? Can we write down all the natural numbers reaching 10^{100}? If not, then according to Aristotle's terminology, one can consider 10^{100} as a number lying 'outside of what has been taken'. This may appear similar to the sense of infinite. On the other hand, theoretically one can precisely define 10^{100} as a number representing 100 times of 10 and hence the number is considered as a finite number. It seems, as if there is a passage of gradually increasing large numbers connecting 'finte' to 'infinite' and 10^{100} is lying in that passage. So, for such a number having inclination towards both 'finite' and 'infinite', can we think of a notion like 'boundary case of infinity'?

Let us consider the phrase that 'there is infinite number of elements in X' as a vague concept. We would like to introduce a quantifier viz., $\exists_{inf}$, 'there exists infinitely many' in the language of $\mathcal{L}_{\mathcal{R}}$. Then a sentence of the form '$\exists_{inf} v(v \in \overline{X})$' can be formed and $Inf(\overline{X})$, (X is infinite) may stand for the expression $\exists_{inf} v(v \in \overline{X})$. Now, interpreting $\exists_{inf}$ as a fuzzy subset assigning a value from $[0, 1]$ to each subset of R, to some extent, we can be precise about a set's grade of infiniteness or in other words, largeness. And then in the context of graded consequence, we can be inquisitive about the classical theorem viz., 'x is a limit point of X if and only if every neighbourhood of x contains infinitely many points of X'. From the perspective of graded consequence, now the concern may extend to study the grade of $\{Lim(\overline{x}, \overline{X})\} \mid\approx \forall u[\exists_{inf} v(v \in N_u(\overline{x}) \& v \in \overline{X})]$, where both $Lim(\overline{x}, \overline{X})$ and $\forall u[\exists_{inf} v(v \in N_u(\overline{x}) \& v \in \overline{X})]$ are treated equivalently in the classical context. Naturally, some well known results like, Bolzano-Weierstrass theorem which states that 'every bounded, infinite set of real numbers has a limit point in R' needs to be investigated in the graded context, where the grade of $\{Inf(\overline{X}), Bdd(\overline{X})\} \mid\approx \exists v Lim(v, \overline{X})$ may not be the same as the grade of $\{Inf(\overline{X}), Bdd(\overline{X})\} \mid\approx \forall u[\exists_{inf} v(v \in N_u(\overline{x}) \& v \in \overline{X})]$.

Classical theory of Mathematics is well established with its notion of infinity. Besides, the concept of infinity also further had been classified in a hierarchical manner. In this initial stage of graded mathematics we are introducing 'there exists infinitely many' as a vague concept. Further study on graded mathematics may throw light on the other related issues.

In 2006, on the occasion of felicitating Prof. Mihir K. Chakraborty on his 60th birthday, an international seminar was organized by Calcutta Logic Circle. The seminar ended with Prof. Chakraborty's talk briefing his area of interests, works and the areas where he desires to work. Graded Mathematics has been one such area which he mentioned as a task to be done. Afterwards, on a few occasions, I had some interactions with him on the idea of graded Mathematics. Here an attempt to develop that idea of graded Mathematics has been made, whereby I took the privilege of giving a shape to our teacher's wish.

References

1. Bandegem Jean Paul Van: Alternative Mathematics: the vague way. Synthese 125: Kluwer Academic Publishers, Netherlands, 19-31, (2000).
2. Chakraborty M.K.: Use of fuzzy set theory in introducing graded consequence in multiple valued logic. in M.M. Gupta and T. Yamakawa (eds.), Fuzzy Logic in Knowledge-Based Systems, Decision and Control, Elsevier Science Publishers, B.V.(North Holland) 247-257, (1988).
3. Chakraborty M.K.: Graded Consequence: further studies. Journal of Applied Non-Classical Logics, 5(2), 127-137, (1995).
4. Goguen J.A.: L-fuzzy Sets. Journal of Mathematical Analysis and Applications, 18(1), 145-174, (1967)
5. Ha$\acute{j}$ek P.: Metamathematics of Fuzzy Logic. Kluwer Academic Publishers, (1998).
6. Lear Jonathan: Aristotelian Infinity. Proceedings of Aristotelian Society, New series, Blackwell Publishing, Vol 80, 187-210, (1979-1980).
7. Pavelka J.: On fuzzy logic I, II, III. Zeitscher for Math. Logik und Grundlagen d. Math 25, 45-52, 119-134, 447-464, (1979).
8. Novak V.: On syntactico-semantical completeness of first order fuzzy logic, Parts I And II. Kybernetica, 2, 6(1,2), Academia Praha, 47-154, (1990).
9. Zadeh L.A.: Fuzzy sets. Information and Control, 8, 338-353, (1965).

Soma Dutta
Indian Institute of Bio-Social Research and Development,
Prafulla Kanan, VIP Road, Kestopur, Kolkata.
somadutta9@gmail.com

Belief and disbelief: a logic approach

Sujata Ghosh

Abstract. Over the years, several researchers have considered giving negative epistemic attitudes such as disbelief a similar footing as the usual ones such as knowledge and belief. This paper provides a survey of such works which focus on the frameworks of belief and disbelief and their interplay. We also provide a few pointers towards further research regarding the dynamics of the removal of inconsistent attitudes that a human mind might possess in certain situations.

1 Introduction

Writing an article for this volume encourages us to be as non-classical and inconsistent as possible, as the main philosophical stance that we have learnt from Mihir Chakraborty is to be tolerant towards different styles and methodologies, both classical and non-classical. Over the years, there has always been an inherent interest towards the so-called non-classical approaches such as fuzzy set theory, rough set theory, vagueness, and non-monotonic logics in developing formal frameworks for reasoning in different scenarios. Akin to these approaches, various formal frameworks have been developed to put forward the notion of *agents' disbeliefs* as a separate epistemic category, in addition to the well-studied notion of *agents' beliefs*. This paper attempts to provide a descriptive summary of some of these developments together with a comparative study. Focussing on the ideas developed in [1,2,3,4,5,6], we look into the interplay of agents' beliefs and disbeliefs.

Considering disbelief together with belief leads us to express various notions of uncertainties as well. For example, having both belief and disbelief over a proposition can be thought of as being uncertain about that proposition, whereas having neither can be thought of as having no opinion. The sky may not be dark with thunder clouds hovering, but it may not be sunny either, leading to uncertainty in an agent's mind regarding the possibility of rain in a few hours time. This mental state would be quite different of having no opinion about the performance of the home team in a cricket match, when one is simply not interested in the game of cricket.

Before proceeding any further with these issues, we would like to take this opportunity to make a thematic translation of an excerpt from a poem written by Mihir Chakraborty [7] to provide an exemplar for his thoughts on the notions of belief and disbelief and their interplay.

The twin rivulets of belief and disbelief
originating from the same bedrock,

to the quiet deluge of which
time and again we retreat,
to descant later.

Flowing alongside,
along their paths, obeying their maxims;
the burden on belief – a wee bit high,
frown it might, under the strain.
On disbelief too, the burden enforced,
to substantiate its presence;
creating a perpetual whirl, even fine-tuned.

I care not, this ostensible difference
in my eyes, they are alike – captivating, alluring
with identical morale and candor.

But, what if they blend and cajole together,
that inconsistent look;
can this world endure?
A patient wait for the new cells to develop,
bringing in new life;
bifurcation happens, creating an ease.

The journey continues,
diverging forces moving into a confluence,
disembodies again.
Unwanted the burden is not,
revisiting the mother rock.

Let us now try to decipher these lines in our way: There is a continuous reasoning process going on in our mind, which results in the formation of various beliefs and/or disbeliefs. We have these beliefs/disbeliefs about the world around, which are more often than not, distinct from each other. But sometimes, it may so happen that our belief and disbelief coincide, giving rise to an inconsistent attitude. What do we do, when something like that happens? If we reach at some kind of inconsistency, our natural tendency would be to try to remove such inconsistencies. We might try to go back in our reasoning process, find the root cause of the problem, and revise some of our assumptions accordingly, so as to get rid of the inconsistency.

Following the lines described above, a formal framework was developed in [4], which we will present in the next section. In addition, we will describe various other frameworks that consider belief and disbelief as separate epistemic attitudes and not as one attitude being the negation of the other. An obvious question arises: Are there situations where it is meaningful to think about belief and disbelief as separate entities? Suppose we believe that a horse will win a race if its rating is above a certain number, and disbelieve it if it is below some other number. *Disbelieving* will not equate to *not believing* in this case, nor to *believing that not*.

Let us consider a fair 100 ticket lottery. Though we do believe that exactly one of the tickets will win the lottery, we have doubts regarding the possibility of an individual ticket winning it. This is the *lottery paradox*, first considered by Kyburg [8]. For some interesting philosophical treatments of this problem, see [9,10]. Another way of resolving this paradoxical situation is by replacing the classical negation (believing that the ticket number 99 will not win) by the weaker notion of disbelief (disbelieving that the ticket number 99 will win). This applies to many practical situations as well. A crime has been committed and two of your very good friends are the prime suspects. It is really hard for you to believe that any one of them has committed the crime, yet the circumstantial evidence forces you to believe that either of them did it.

From the perspective of belief merging, consider k sources of information providing their opinions regarding a certain event p. Suppose that m of them state that p holds, and n of them state that p does not hold. Some of the sources may not have any opinion regarding p, but none of the sources are inconsistent in the sense that any single source does not simultaneously state that p holds and does not hold. So we have that $m + n \leq k$. The fraction m/k can be seen as the degree of certainty of the source that p holds and n/k that ¬p holds. Let $cr(p) \in [0, 1]$ denote the degree of certainty that p holds. We can think of threshold values t_1 and t_2 $(0 < t_1 \leq t_2 \leq 1)$ for belief and disbelief, that is, p is believed if $t_2 \leq cr(p)$, and is disbelieved if $cr(p) < t_1$. In the remaining cases p is neither believed nor disbelieved. So, from the fact the p is not believed, we cannot say that p is disbelieved. Also, from the fact that p is disbelieved, it does not necessarily follow that $\neg p$ is believed.[1]

Considering belief and disbelief together is in some sense talking about both positive and negative attitudes of a person. The idea of representing both positive and negative aspects of a subject is not new. There are approaches with such proposals in many areas, like decision theory [12], argumentation theory [13], and many others (see [14] for an overview). The concept of *bipolarity* is precisely about this: an explicit handling of the positive and negative aspects in information [15]. It is based on the fact that, when taking a decision or weighing some possibilities, we consider not only the positive aspects of the available options, but also the negative ones.

From this perspective, frameworks that consider only the positive aspect can be seen as special situations in which the positive and the negative information are mutually exclusive and mirror images of each other: I consider p as good if and only if I consider $\neg p$ as bad. But this does not need to be the case: we can imagine a situation in which, though p is good, its negation $\neg p$ is not necessarily bad, and the notion of bipolarity allows us to deal with such cases. Considering such dual frameworks for positive and negative information paves the way for an in-depth study of qualitative representation of uncertainty [6].

[1] These examples are taken from [3,11].

In the remaining part of this paper, we provide a comprehensive survey of some of the research done in this area in Section 2. A technical comparison for a few frameworks is provided in Section 3. Section 4 suggests a dynamic framework to model the removal of inconsistent attitudes. Pointers towards future work have been included in Section 5.

2 A descriptive survey

We now discuss various formal frameworks that have evolved over the years providing different ways for modeling the notion of disbelief, focussing on the relevant aspects of such models. We start with describing Gomolinska's work on developing a logic of acceptance and rejection [1], followed by a short survey of several logics of belief and disbelief proposed by Chopra et al [3]. Then we move on to discuss a procedural framework of revising beliefs and disbeliefs [4], and finally finishing with discussing more recent works on modeling disbelief [5,6]. The main objective of this survey is to acquaint the reader with the varied formalisms developed to deal with the notion of *disbelief of an agent* explicitly. A technical comparison of some of these works will be provided in Section 3.

2.1 A logic of acceptance and rejection

Gomolinska [1] presented a nonmonotonic formalism, AEL2, extending the framework of Moore's autoepistemic logic [16] to deal with *uncertainty* of an agent. In order to deal with uncertainty that arises in an agent's mind, the possible alternatives regarding a piece of information that are taken into account are that of acceptance, rejection, and lack of decision. Even though the underlying intuition came from a different motivation, that of modeling uncertainty, it was felt that a more fine-grained study of what can be accepted (believed) or rejected (disbelieved) was needed. We now give a short formal account of this study.

The underlying logical framework is that of classical propositional logic augmented with two modal operators, B and D, for acceptance and rejection, respectively, where B and D are not inter-definable in general. In AEL2, accepted and rejected premises are separated to form a pair of sets of formulas, say (I_1, I_2), and AEL2 extensions (T_1, T_2) of (I_1, I_2) are considered, where T_1 is expected to contain all the accepted formulas with respect to I_1 and T_2 to contain all the rejected formulas with respect to I_2. For (T_1, T_2) to be considered as stable expansions of (I_1, I_2), T_1 needs to be closed under the classical consequence operator Cn, and T_2 needs to be closed under the rejection consequence operator Cn'. To define these operators, let us first define what we mean by a syntactic consequence operator C [1].

Consider an inference rule r. The notation $(Y, \alpha) \in r$ is used to indicate that Y is a finite set of premises and α is a conclusion of an inference rule r from the premise Y. Given a set of inference rules R and a set of formulas X, C is defined inductively as follows:

$C^0(R, X) = X$;
$C^{n+1}(R, X) = C^n(R, X) \cup \{\alpha \mid \exists_{r \in R} \exists_{Y \subseteq C^n(R,X)} (Y, \alpha) \in r\}$;
$C(R, X) = \bigcup_{n \in \mathbb{N}} C^n(R, X)$.

Let us consider the inference rules MP (modus ponens) and Rej, given by:

$$MP : \frac{\alpha, \alpha \to \beta}{\beta} \qquad\qquad Rej : \frac{\beta, \neg(\alpha \to \beta)}{\alpha}$$

Note that $\neg(\alpha \to \beta)$ is equivalent to $\alpha \wedge \neg\beta$. Intuitively, the Rej rule says that if both β and $\alpha \wedge \neg\beta$ do not hold, then it follows that α cannot hold. Let $\mathcal{L}_{BD}$ denote the language of AEL2, and FOR_{BD} denote the set of all formulas in $\mathcal{L}_{BD}$. Let Ax denote the set of all substitution instances of the classical propositional logic axioms in the language of $\mathcal{L}_{BD}$. Then for a set of formulas $X \subseteq \mathsf{FOR}_{BD}$, $Cn(X)$, and $Cn'(X)$ are defined as follows:

$$Cn(X) = C(\{MP\}, \mathsf{Ax} \cup X)$$
$$Cn'(x) = C(\{Rej\}, \neg Cn(\emptyset) \cup X), \text{ where } \neg Y = \{\neg\alpha : \alpha \in Y\}$$

We say that a set of formulas X is Cn-consistent if $Cn(X) \neq \mathsf{FOR}_{BD}$, Cn'-consistent if $Cn'(X) \neq \mathsf{FOR}_{BD}$. A pair of sets of formulas in $\mathcal{L}_{BD}$, (X, Y) is said to be consistent if X is Cn-consistent, and Y is Cn'-consistent. Now we are all set to define stable expansions of a pair of accepted and rejected sets of formulas due to [1].

Definition 2.1. (T_1, T_2) *is said to be a stable* AEL2 *expansion of* (I_1, I_2) *if,*
$T_1 = Cn(I_1 \cup \{B\varphi : \varphi \in T_1\} \cup \{\neg B\varphi : \varphi \notin T_1\} \cup \{D\varphi : \varphi \in T_2\} \cup \{\neg D\varphi : \varphi \notin T_2\})$*;*
$T_2 = Cn'(I_2 \cup \{\neg B\varphi : \varphi \in T_1\} \cup \{B\varphi : \varphi \notin T_1\} \cup \{\neg D\varphi : \varphi \in T_2\} \cup \{D\varphi : \varphi \notin T_2\})$.

The above definition gives a syntactic way of producing stable sets of accepted and rejected propositions, with respect to some initial configurations. The intuitive meaning of such sets being stable is that : a) no further conclusions can be drawn from them, and b) a set of accepted propositions should be in some sense grounded in terms of the initially accepted propositions, and the same for a set of rejected propositions. In terms of epistemic attitudes, these stable expansions can be thought of as *intuitively complete* pairs of sets of beliefs and disbeliefs that an agent might have, starting from her initial position. Gomolinska has provided semantic characterizations of these notions as well [1]. Based on this logical framework, Gomolinska and Pearce [11] developed a formalism for disbelief expansion, contraction and revision in line with the corresponding Alchourrón-Gärdenfors-Makinson (AGM) belief operations [17].

2.2 Some logics of belief and disbelief

Intuitively, there is a lot of interaction that goes on between believing and disbelieving propositions and negations of such propositions. That is essentially the reason to consider belief and disbelief as separate epistemic entities. In [3], Chopra, Heidema and Meyer presented four logics to describe agents' beliefs and disbeliefs and their interplay. These formalisms differ in terms of their usage of the negation, from classical to weaker ones, as well as in terms of the interrelationship assumed between beliefs and disbeliefs. Below we will first present the general framework and then the four logics presented in [3], namely, i) a weak logic of belief and disbelief, **WBD**, ii) a logic of belief and disbelief similar to that developed by Gomolinska, **GBD**, iii) a logic describing some interactions between belief and disbelief, **BD**, and finally, iv) a logic where disbelief corresponds to classical negation, **BN**.

The framework Given a set of atomic propositions P, let L_B denote the set of propositional formulas built from P, and let $L_D := \{\overline{\phi} \mid \phi \in L_B\}$. The language L is given by $L_B \cup L_D$. The agent's information set, containing her beliefs and disbeliefs, is defined as any subset of the language. Given an information set Γ,

- the agent is said to *believe* every ϕ such that $\phi \in \Gamma$ (forming $\Gamma_B \subseteq \Gamma$) , and
- the agent is said to *disbelieve* every ϕ such that $\overline{\phi} \in \Gamma$ (forming $\Gamma_D \subseteq \Gamma$).

Based on this language and the corresponding representation of an agent's beliefs and disbeliefs, the authors propose four different logics by defining different closure properties an agent's information set should have. We are only talking about the single agent case here, but one can easily extend these logics to describe the multi-agent systems.

The logic WBD This logic is defined by the following three properties of its consequence relation $\vdash$:

(B) If ϕ is a propositional consequence of what the agent believes, then the agent believes it:
$$\Gamma_B \vdash_{PL} \phi \text{ implies } \Gamma \vdash \phi$$

(D$\perp$) If ϕ is a propositional contradiction, then the agent disbelieves it:
$$\phi \vdash_{PL} \perp \text{ implies } \Gamma \vdash \overline{\phi}$$

(WD) If the agent disbelieves ψ and ϕ implies ψ classically, then the agent disbelieves ϕ:
$$\overline{\psi} \in \Gamma_D \text{ and } \phi \vdash_{PL} \psi \text{ imply } \Gamma \vdash \overline{\phi}$$

Note that in this framework we do not find any interrelation between the derivation processes of beliefs and disbeliefs.

Let us now have a look at the semantics. Let V denote the set of all propositional valuations for the atomic propositions in the language. A **WBD** model is a tuple $\langle M, \mathcal{N} \rangle$ where $M \subseteq \mathsf{V}$ is a set of propositional valuations, that is, a set of possible worlds with valuation already provided, and $\mathcal{N} \subseteq \wp(V)$ is a set of sets of propositional valuations. Intuitively, M represents the positive information, that is, the situations the agent considers possible, and $\mathcal{N}$ is a collection of an agent's sources which gives the information regarding what can be disbelieved. Each $N \in \mathcal{N}$ is a set of the situations ($N \subseteq \mathsf{V}$) which are considered possible by such a source. An agent is associated with at least one source of negative information, that is, $\mathcal{N} \neq \emptyset$. Given a **WBD** model,

$$\langle M, \mathcal{N} \rangle \models \phi \quad \text{iff} \quad \phi \text{ is true under every valuation in } M \text{, and}$$
$$\langle M, \mathcal{N} \rangle \models \overline{\phi} \quad \text{iff} \quad \neg\phi \text{ is true under every valuation of } N \textit{ for some } N \in \mathcal{N}.$$

The agent believes ϕ ($\phi \in L$) whenever $\langle M, \mathcal{N} \rangle \models \phi$, and the agent disbelieves ϕ ($\overline{\phi} \in L$) whenever $\langle M, \mathcal{N} \rangle \models \overline{\phi}$. The consequence relation $\vdash$ characterized by the three properties **B**, **D**$\perp$ and **WD** mentioned above corresponds exactly to the semantic consequence relation $\models$ just defined.

The logic GBD The logic **GBD** considers a consequence relation for disbeliefs which is exactly the dual of that for beliefs (similar to the approach in [1,11]).

Let $\overline{\Gamma}_D$ denote the set containing the negation of the propositional formulas that correspond to formulas in Γ_D, that is,

$$\overline{\Gamma}_D := \{\neg\phi \mid \overline{\phi} \in \Gamma_D\}$$

In other words, $\overline{\Gamma}_D$ contains the negations of the formulas the agent disbelieves. Then consider the following property:

(GD) If $\neg\phi$ can be concluded from $\overline{\Gamma}_D$, the agent disbelieves ϕ:

$$\overline{\Gamma}_D \vdash_{PL} \neg\phi \;\text{ implies }\; \Gamma \vdash \overline{\phi}$$

This condition exemplifies the role of negation in considering disbeliefs of an agent with belief playing an implicit role. The logic **GBD** is obtained from **WBD** by replacing **WD** by **GD**.

The class of **WBD** models that corresponds to the **GBD** logic contains exactly those models $\langle M, \mathcal{N} \rangle$ in which $\mathcal{N}$ has exactly one element. This assumption is quite natural: the **B** property states that the set of an agent's beliefs is closed under logical consequences, but the **GD** property states that the set of an agent's disbeliefs is closed under the logical consequence modulo negation of formulas.

Note again that the relation between beliefs and disbeliefs is still at a superficial level; the closure properties for the disbelief notion have only been strengthened in comparison to **WBD**. One interesting observation about this logic is the following: if the agent disbelieves both ϕ and ψ then she disbelieves $\phi \vee \psi$. But then, this notion of disbelief is on similar footing as the classical notion of negation, which is not desirable. Returning to the example from the introduction, it may be very hard to believe that your friend Craig is the traitor and even that another close friend Denis is the traitor, but circumstantial evidence may make it perfectly plausible that one of them is. The subsequent logic, **BD**, rectifies this disadvantage.

The logic BD The logic **BD** provides some interaction between beliefs and disbeliefs, in the sense that *disbelieving* is considered to be a weaker notion than *believing not*. Formally, **BD** is defined by **B**, **D**$\bot$, and the following property:

(D) If the agent disbelieves some ψ which is a classical consequence of the agent's beliefs together with ϕ, then the agent disbelieves ϕ:

$$\Gamma \vdash \overline{\psi} \text{ and } \Gamma_B \cup \{\phi\} \vdash_{PL} \psi \text{ imply } \Gamma \vdash \overline{\phi}$$

The general idea is that one should disbelieve those propositions whose consequences are disbelieved. The notion of consequence can be independent of other factors (**WD**), or may be dependent or grounded in the agent's belief set (**D**). In some sense, property **WD** is the special case of **D** where $\Gamma_B = \emptyset$.

A consequence of this property: For any ϕ, believing in $\neg\phi$ implies disbelieving ϕ. For, by property **D**$\bot$, the agent disbelieves any contradiction, that is, $\Gamma_D \vdash \overline{\phi \wedge \neg\phi}$. Suppose the agent believes in $\neg\phi$; since $\{\neg\phi, \phi\}$ implies classically that $\phi \wedge \neg\phi$, that is, $\{\neg\phi, \phi\} \vdash_{PL} \phi \wedge \neg\phi$, the agent disbelieves in ϕ, as desired.

The class of **WBD** models that correspond to the **BD** logic contains exactly those models $\langle M, \mathcal{N} \rangle$ in which every $N \in \mathcal{N}$ is a subset of M. This is to say that we get the logic **BD** when the agent already considers possible all the situations that each one of her sources considers possible.

The fact that the agent can disbelieve both ϕ and $\neg\phi$ without disbelieving $\phi \vee \neg\phi$, follows from the fact that now the agent can have more than one source to acquire her disbeliefs.

The logic BN Finally, the logic **BN** looks at the other direction, adding a property to **BD** to guarantee that if the agent disbelieves a formula, then she believes in its negation. This is given by the following property:

(D $\rightarrow$ B) If the agent disbelieves ϕ, then she believes in $\neg\phi$:

$$\Gamma \vdash \overline{\phi} \text{ implies } \Gamma \vdash \neg\phi$$

Because the logic **BD** already satisfies the other direction, adding this property to **BD** gives us: *The agent disbelieves ϕ if and only if she believes in $\neg\phi$.* Then, the notion of disbelief collapses to classical negation. For the semantics, $\langle M, \mathcal{N} \rangle$ are those models such that $\mathcal{N}$ is a singleton and consists of M.

Some discussions It is evident that there can be various ways of constructing formal syntax and semantics to reason about agents' beliefs and disbeliefs. But the interplay between an agent's belief and disbelief needs to be taken into account when one is modeling such notions. Otherwise, one could very well propose certain technical results without any intuitive grounding. In this sense, the logics **BD** and **BN** are more interesting as they talk about generating disbeliefs from beliefs and beliefs from disbeliefs. But, as Chopra and colleagues [3] mention, we do not really get past the expressivity of propositional logic in this whole exercise. In what follows, we introduce several other ways of describing the notion of disbelief that do transport us beyond propositional logic, and provide pointers to the readers regarding the usefulness of such methodologies.

2.3 A procedural framework for revising belief and disbelief

Influenced by the work on disbelief change [11], Ghosh [4] developed a procedural framework for revising or updating beliefs and disbeliefs in a step-by-step manner described below. The whole framework is based on a belief-disbelief duo, a pair of sets $\langle B, D\rangle$ representing belief and disbelief. This study proposes a general method for deriving a pair $\langle \beta, \delta\rangle$ of sentences from a pair of sets $\langle B, D\rangle$. In addition, two operations, namely *deletion* of a formula and *replacement* of a formula by another, are performed on $\langle B, D\rangle$ giving rise to a new pair, $\langle B', D'\rangle$. These operations correspond to *AGM-contraction* and *AGM-revision* respectively [17] and are performed only when the *need* arises. The need informally means that at least one statement has been *derived* as belief and disbelief both. There are some restrictions on replacement, that is, not any sentence is allowed to replace an existing one. This means that a sentence (belief/disbelief) may be replaced by any of a specified collection of sentences, unlike for standard revisions.

Consider an arbitrary propositional language L and a suitable algebraic structure for its interpretation (i.e. semantics). A valuation v is a homomorphism from L to the corresponding algebraic structure. Two (possibly different) logical frameworks are considered together in a pair form which is termed as a *bi-logic*.

A bi-logic $\Vdash$ is a consequence relation deriving a pair $\langle \beta, \delta\rangle \in L \times L$ from a pair $\langle B, D\rangle \in \mathcal{P}(L) \times \mathcal{P}(L)$, denoted by, $\langle B, D\rangle \Vdash \langle \beta, \delta\rangle$, and defined by two consequence relations $\vdash_1$ and $\vdash_2$ from $\mathcal{P}(L)$ to L, such that $B \vdash_1 \beta$ and $D \vdash_2 \delta$, where both of $\vdash_1$ and $\vdash_2$ may be inter-defined or independent. In any case, each of the consequences, $\vdash_1$ and $\vdash_2$ is called the dual of the other and each of the premise sets B, D is called the dual of the other.

Two subsets D_1 and D_2 (with $D_1 \cap D_2 = \emptyset$) of the base set of the algebraic structure are taken as designated sets for the two logics, given by $\vdash_1$ and $\vdash_2$ respectively. Relative to D_1 and D_2, semantic consequence relations $\models_1$ and $\models_2$ are defined. It is assumed that $X \vdash_1 \alpha$ iff $X \models_1 \alpha$ and $X \vdash_2 \alpha$ iff $X \models_2 \alpha$, i.e. the logics are sound and complete relative to the respective semantics. For simplicity, we assume that both the base logics have Hilbert-style axiomatization.

Let us now describe the operations in more detail. To take care of the restricted replacement operation, two binary relations R_1 and R_2 are defined in the language that satisfy the following conditions:

$$\text{If } \alpha\ R_1\ \beta \text{ and } \beta \vdash_1 \gamma, \text{ then } \alpha \vdash_1 \gamma$$
$$\text{If } \alpha\ R_2\ \beta \text{ and } \beta \vdash_2 \gamma, \text{ then } \alpha \vdash_2 \gamma$$

Let $X \subseteq$ L and $\alpha \in L$. Let $X \vdash_i \alpha$, i = 1,2 and let a particular derivation of α be specified. Let $P^X_{r_i}(\alpha)$ denote a minimal subset X' of X such that the given derivation is still a derivation of α from X'.This is a finite subset of X. We impose the additional restriction that X' does not include any axiom.

Now, if $X \models_i \alpha$, i.e. α semantically follows from X, then because of completeness, $X \vdash_i \alpha$ and some $P^X_{r_i}(\alpha)$ exists. The notation $X(\alpha/\beta)$ denotes the set of formulas obtained from X by replacing the formula α by β; the set $X(\alpha/\beta)$ is said to be obtained from X by replacement.

A logic for belief and disbelief $\mathcal{L}$ is defined as follows. Let $\langle B, D\rangle$ be a pair such that B, $D \subseteq L$ are consistent subsets relative to the respective logics.

Definition 2.2. *The syntactic consequence relation $\Vdash$ of the logic $\mathcal{L}$ shall be defined through the following steps.*
(1) A derivation of the pair $\langle \beta, \delta\rangle$, β, $\delta \in L$ in one step by rules in RS_1 and RS_2 denoted by $\Rightarrow$ is defined as follows:

$$\langle B, D\rangle \Rightarrow \langle \beta, \delta\rangle \text{ iff either } \delta \in D \text{ and } B \vdash_1 \beta \text{ in one step;}$$
$$\text{or, } \beta \in B, \text{ and } D \vdash_2 \delta \text{ in one step,}$$

where $A \vdash_i \alpha$ iff either $\alpha \in A$, or α is an axiom or obtained by some rule. Such an expression $\langle B, D\rangle \Rightarrow \langle \beta, \delta\rangle$ is called a sequent.
(2) $\langle B, D\rangle \Vdash \langle \beta, \delta\rangle$ iff there exists a sequence of sequents,

$$\langle B_i, D_i\rangle \Rightarrow \langle \beta_i, \delta_i\rangle,\ i = 0,1,\ \ldots,\ n$$

to some of which a $*$–mark is ascribed satisfying the following conditions :
(1a) $\langle B_0, D_0\rangle = \langle B, D\rangle$.
(1b) $\langle \beta_n, \delta_n\rangle = \langle \beta, \delta\rangle$.
(2) $\langle B_0, D_0\rangle \Rightarrow \langle \beta_0, \delta_0\rangle$ is $*$–marked.
(3) If $\langle B_{i+1}, D_{i+1}\rangle \Rightarrow \langle \beta_{i+1}, \delta_{i+1}\rangle$ is $*$–marked then, $(B_i \cup \{\beta_i\}) \cap (D_i \cup \{\delta_i\}) \neq \emptyset$.
(4) either, $B_{i+1} = B_i \cup \{\beta_i\}$, $D_{i+1} = D_i \cup \{\delta_i\}$,
or, $\langle B_{i+1}, D_{i+1}\rangle$ is obtained by DELETION, i.e.

$$B_{i+1} = B_{f(i)} - \{\beta'\} \text{ for some } \beta' \in P_{r_1}^{B_{f(i)}}(\alpha) \text{ and } D_{i+1} = D_{f(i)}$$

or,

$$D_{i+1} = D_{f(i)} - \{\delta'\} \text{ for some } \delta' \in P_{r_2}^{D_{f(i)}}(\alpha) \text{ and } B_{i+1} = B_{f(i)},$$

or, $\langle B_{i+1}, D_{i+1}\rangle$ is obtained by REPLACEMENT, i.e.

$$B_{i+1} = B_{f(i)}(\beta''/\beta') \text{ s.t. } \beta''\ R_1\ \beta' \text{ for some } \beta' \in P_{r_1}^{B_{f(i)}}(\alpha) \text{ and } D_{i+1} = D_{f(i)}$$

or,

$$D_{i+1} = D_{f(i)}(\delta''/\delta') \text{ s.t. } \delta''\ R_2\ \delta' \text{ for some } \delta' \in P_{r_2}^{D_{f(i)}}(\alpha) \text{ and } B_{i+1} = B_{f(i)},$$

where $\alpha \in (B_i \cup \{\beta_i\}) \cap (D_i \cup \{\delta_i\})$, and f(i) is the largest integer $\leq i$ such that the f(i)-th sequent is $*$–marked. This $*$–mark is used to denote some special steps in the derivation, where the possibilities of deletion or replacement may come up.

A typical derivation with n = 11, takes the form:

$$\begin{array}{l} * \langle B_0, D_0\rangle \Rightarrow \langle \beta_0, \delta_0\rangle \\ \ \ \langle B_1, D_1\rangle \Rightarrow \langle \beta_1, \delta_1\rangle \\ \ \ \langle B_2, D_2\rangle \Rightarrow \langle \beta_2, \delta_2\rangle \end{array}$$

$$
\begin{aligned}
* \; & \langle B_3, D_3 \rangle \Rightarrow \langle \beta_3, \delta_3 \rangle \\
& \langle B_4, D_4 \rangle \Rightarrow \langle \beta_4, \delta_4 \rangle \\
* \; & \langle B_5, D_5 \rangle \Rightarrow \langle \beta_5, \delta_5 \rangle \\
& \langle B_6, D_6 \rangle \Rightarrow \langle \beta_6, \delta_6 \rangle \\
& \langle B_7, D_7 \rangle \Rightarrow \langle \beta_7, \delta_7 \rangle \\
* \; & \langle B_8, D_8 \rangle \Rightarrow \langle \beta_8, \delta_8 \rangle \\
* \; & \langle B_9, D_9 \rangle \Rightarrow \langle \beta_9, \delta_9 \rangle \\
& \langle B_{10}, D_{10} \rangle \Rightarrow \langle \beta_{10}, \delta_{10} \rangle \\
& \langle B_{11}, D_{11} \rangle \Rightarrow \langle \beta_{11}, \delta_{11} \rangle
\end{aligned}
$$

At the $*$ – marked steps 3, 5, 8, 9, either a deletion or a replacement is performed. In the above derivation, $f(2) = 0, f(4) = 3, f(7) = 5, f(8) = 8$.

The sequence $\langle B_i, D_i \rangle$ denotes the successive belief-disbelief pairs of sets (called b-d states). The pair $\langle \beta_i, \delta_i \rangle$ denotes the belief-disbelief pair of formulas obtained at the i-th stage from $\langle B_i, D_i \rangle$ by the base logics by one-step derivation. Any such $\langle B_i, D_i \rangle$ is also said to be *derived* from the initial b-d state $\langle B, D \rangle$.

The semantic relation for this logic can be defined in a similar fashion as the syntactic one [4]. The derivation process describes the step-by-step reasoning method that an agent performs while updating her beliefs and disbeliefs about the world around. The interesting part is when the agent has some inconsistencies in her mind. A procedure for removing such inconsistencies has been provided above. These notions are formalized below.

Definition 2.3. *A pair $\langle B, D \rangle$ is said to be incoherent iff there exists a b-d state $\langle B_i, D_i \rangle$ derived from $\langle B, D \rangle$, by the consequence relation $\Vdash$, such that $(B_i \cup \{\beta_i\}) \cap (D_i \cup \{\delta_i\}) \neq \emptyset$, where $\langle B_i, D_i \rangle \Rightarrow \langle \beta_i, \delta_i \rangle$ is a sequent in the derivation.*

Definition 2.4. *A pair $\langle B, D \rangle$ is said to be coherent iff it is not incoherent.*

It can be shown that a pair $\langle B, D \rangle$ is incoherent iff $\{\beta : \mathrm{B} \vdash_1 \beta\} \cap \{\delta : \mathrm{D} \vdash_2 \delta\} \neq \emptyset$. An agent's natural state of mind is often incoherent in the above sense, and the target is to reach a coherent state of mind through a reasoning process. This exercise is meant to model such a process. The step-by-step derivation method is meant to capture the dynamics of an agent's reasoning. We will come back to this discussion on the dynamic removal of agents' inconsistencies in Section 4.

2.4 Modeling disbelief in terms of belief ordering

De Jongh and Ghosh [5] used the notion of belief ordering to describe disbeliefs. The authors gave different status to belief and disbelief, considering belief to be a more fundamental notion in lines of the **BD**-logic described in [3]. Describing belief ordering explicitly provides this path of expressing the concept of disbelief. In this sense, disbelieving a proposition means that the inclination to believe in its negation is stronger than the inclination to believe it. Consequently, disbelieving is a much weaker notion than believing the negation of the proposition, but on the other hand, it should imply that one does not believe in the proposition. In other words, $D\varphi$ is implied by $B\neg\varphi$ and implies $\neg B\varphi$ but not the other way around in either case.

Consider the following exemplification of such a situation. Due to the unpredictable weather conditions, Pam's belief that it is not a good idea for her to bike from Amsterdam to Leiden is much stronger than her belief that it is a good idea. When options like these are available, it is very natural to have this sort of ordering dilemma playing around people's mind. This can be interpreted as that Pam disbelieves that it is a good idea to bike, which evidently implies that she does not believe that it is a good idea to bike. But *she believes that it is not a good idea to bike* is a much stronger statement, which fails to express the finer interplay of doubts that may be prevalent in one's mind.

Before going any further, let us first introduce the framework of belief ordering. The language is defined as follows:

Definition 2.5. *Given a countable set of atomic propositions Φ, formulas φ are defined by the Backus-Naur form:*

$$\varphi := \bot \mid p \mid \neg\varphi \mid \varphi \vee \varphi \mid B\varphi \mid \varphi \succcurlyeq_B \psi$$

where $p \in \Phi$.

The intuitive reading of the formula $B\varphi$ is φ *is believed*, and that of $\varphi \succcurlyeq_B \psi$ is *belief in φ is at least as strong as belief in ψ*. We introduce the notations $\varphi \succ_B \psi$ for $(\varphi \succcurlyeq_B \psi) \wedge \neg(\psi \succcurlyeq_B \varphi)$ and $\varphi \equiv_B \psi$ for $(\varphi \succcurlyeq_B \psi) \wedge (\psi \succcurlyeq_B \varphi)$. Intuitively, they can be read as *belief in φ is stronger than that in ψ* and *belief in φ and ψ are of same strength*, respectively. We now move on to define a model for this logic.

Definition 2.6. *A model is defined to be a structure $\mathcal{M} = (S, \leq, \geq_B, V)$, where S is a non-empty finite set of states, V is a valuation assigning truth values to atomic propositions in states, $\leq$ is a quasi-linear*[2] *order relation (a plausibility ordering) over S, and $\geq_B$ is a quasi-linear order relation over $\mathcal{P}(S)$, satisfying the conditions*

1. *If $X \subseteq Y$, then $Y \geq_B X$*
2. *If $\mathcal{B}$ is the set of all $\leq$-minimal worlds (the set of most-plausible worlds, called the* center*), then $\mathcal{B} \subseteq X$ and $\mathcal{B} \not\subseteq Y$ imply $X >_B Y$, where $X >_B Y$ iff $X \geq_B Y$ and not ($Y \geq_B X$).*
3. *If X is non-empty, then $X >_B \emptyset$.*

The truth definition for formulas φ in a model $\mathcal{M}$ is as usual with the following clauses for the belief and ordering modalities.

[2] A binary relation $\leq$ on a non-empty set S is said to be quasi-linear if it is reflexive, transitive and linear, i.e. a total pre-order. That we do take the order to be quasi-linear, but not more generally a pre-order is not a matter of principle but rather of convenience.

$\mathcal{M}, s \models B\varphi$ iff $\mathcal{M}, t \models \varphi$ for all $\leq$-minimal worlds t.

$\mathcal{M}, s \models \varphi \succcurlyeq_B \psi$ iff $\{t \mid \mathcal{M}, t \models \varphi\} \geq_B \{t \mid \mathcal{M}, t \models \psi\}$.

From the definition of $\succ_B$, it follows that:

$\mathcal{M}, s \models \varphi \succ_B \psi$ iff $\{t \mid \mathcal{M}, t \models \varphi\} >_B \{t \mid \mathcal{M}, t \models \psi\}$.

It is clear that $\succcurlyeq_B$ is a global notion: if $\varphi \succcurlyeq_B \psi$ is true anywhere in the model, then it is true everywhere. So, it is either true or false throughout the model; $\succcurlyeq_B$ is a global notion like B. It follows that $\succ_B$ is also a global notion.

Going back to the notion of disbelief as presented in this work, the formal way of representing it is as follows: $D\varphi$ corresponds to $\neg\varphi \succ_B \varphi$: disbelieving φ means that the inclination of believing in $\neg\varphi$ is more than that in φ. Thus, *disbelief* is given a global stance and the representation is different from that of *belief* which is apparent from their respective interpretations. This also emphasizes the fact that disbelieving something is different from both *not believing* as well as from *believing the negation.*

A logic of belief and disbelief was introduced, and the authors provided a complete axiomatization, which is presented below.

Definition 2.7. *Given a countable set of atomic propositions Φ, formulas φ are defined by the Backus-Naur form:*

$\varphi := \bot \mid p \mid \neg\varphi \mid \varphi \vee \varphi \mid B\varphi \mid D\varphi$

where $p \in \Phi$.

Theorem 2.1. *BD-logic is complete and its validities are completely axiomatized by the following axioms and rules:*

a) all substitution instances of propositional tautologies and inference rules
b) all KD45 axioms and rules
c) disbelief axioms:
$D\psi \wedge D\varphi \to D(\psi \vee \neg D\varphi)$
$\neg D\varphi \to DD\varphi$
$D\varphi \to \neg D\neg\varphi$
$D\bot$
d) special axioms:
$B\varphi \to D\neg\varphi$
$D\varphi \to BD\varphi$
e) anti-monotone rule:
if $\varphi \to \psi$ then $D\psi \to D\varphi$.

Here, the *KD45* logic denotes the usual logic representing the belief modality. For more details, see [18].

One can show that the intuitively incorrect principle discussed in Section 2.2, $D\varphi \wedge D\psi \to D(\varphi \vee \psi)$, can be avoided in the BD-logic.

2.5 A modal logic of belief and disbelief

In [6], Ghosh and Veláquez-Quesada proposed a bi-modal framework that allows to express various kinds of attitudes toward a formula φ. Once again, the aim was to deal with negative attitudes of agents on a par with their positive attitudes. An extension of the classical modal language was introduced that allows to express both positive and negative attitudes *explicitly.*

Definition 2.8. *Let* $\mathtt{P}$ *be a set of atomic propositions. Formulas* φ *of the language* $\mathcal{L}$ *are given by*

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid [+]\varphi \mid [-]\varphi$$

with $p \in \mathtt{P}$*. Formulas of the form* $[+]\varphi$ *(*$[-]\varphi$*) are read as the agent has a positive (negative) attitude towards* φ*. The corresponding diamond modalities are defined in the usual way.*

Definition 2.9 (Dual model). *Given a set of atomic propositions* $\mathtt{P}$*, a* dual model *is a tuple* $\mathcal{M} = \langle W, R^+, R^-, V\rangle$ *where* W *is a non-empty set of worlds,* R^+ *and* R^- *are binary relations on* W *and* $V : \mathtt{P} \to \wp(W)$ *is a valuation function. We denote by* $\boldsymbol{M}$ *the class of all semantic models.*

The difference between this system and an ordinary bi-modal framework relies on the interpretation of negative attitude formulas $[-]\varphi$.

Definition 2.10. *Let* $\mathcal{M} = \langle W, R^+, R^-, V\rangle$ *be a dual semantic model and let* w *be a world in it. Atomic propositions, negation and conjunction are interpreted as usual. For the modalities, we have*

$$(\mathcal{M}, w) \models [+]\varphi \quad \textit{iff} \quad \textit{for all } w' \textit{ such that } R^+ww',\ (\mathcal{M}, w') \models \varphi$$
$$(\mathcal{M}, w) \models [-]\varphi \quad \textit{iff} \quad \textit{for all } w' \textit{ such that } R^-ww',\ (\mathcal{M}, w') \models \neg\varphi.$$

We would refer to $[+]$ and $[-]$ as universal modalities, and the corresponding diamond modalities as the existential ones. The authors also considered a particular interpretation of the $[+]$ and the $[-]$ modalities: *belief* and *disbelief.* As mentioned earlier, some intuitive ways to relate them are: *disbelieving* φ is a stronger notion than *not believing in* φ, whereas *believing in* $\neg\varphi$ should imply *disbelieving* φ. In what follows, the model and axiom system of the belief-disbelief logic ($\mathbf{L}_{BD}$) are provided.

Definition 2.11. *Let* $\boldsymbol{M}_{BD}$ *denote the class of models for which the positive and negative relations, now denoted by* R^{B} *and* R^{D}*, are serial, reflexive and Euclidean. Their respective universal modalities are given by* B *and* D *(with* $\widehat{\mathrm{B}}$ *and* $\widehat{\mathrm{D}}$ *denoting the corresponding existential ones).*

Theorem 2.2. *The logic* $\boldsymbol{L}_{BD}$ *given by the axiom system below is sound and complete with respect to* $\boldsymbol{M}_{BD}$*.*

P	*All instances of propositional tautologies*	**MP**	*If* $\vdash \varphi$ *and* $\vdash \varphi \rightarrow \psi$*, then* $\vdash \psi$
K+	$\vdash B(\varphi \rightarrow \psi) \rightarrow (B\varphi \rightarrow B\psi)$	**K-**	$\vdash D(\varphi \wedge \psi) \rightarrow (D\neg\varphi \rightarrow D\psi)$
Gen+	*If* $\vdash \varphi$*, then* $\vdash B\varphi$	**Gen-**	*If* $\vdash \neg\varphi$*, then* $\vdash D\varphi$

D+	$\vdash \mathrm{B}\,\varphi \rightarrow \widehat{\mathrm{B}}\varphi$	**D-**	$\vdash \mathrm{D}\,\varphi \rightarrow \widehat{\mathrm{D}}\varphi$
4+	$\vdash \mathrm{B}\,\varphi \rightarrow \mathrm{B}\,\mathrm{B}\,\varphi$	**4-**	$\vdash \mathrm{D}\,\varphi \rightarrow \mathrm{D}\,\neg\,\mathrm{D}\,\varphi$
5+	$\vdash \neg\,\mathrm{B}\,\varphi \rightarrow \mathrm{B}\,\neg\,\mathrm{B}\,\varphi$	**5-**	$\vdash \neg\,\mathrm{D}\,\varphi \rightarrow \mathrm{D}\,\mathrm{D}\,\varphi$

This can be considered as a minimal logic of belief and disbelief. To make things more interesting and useful one should have inter-relations between the belief and disbelief modalities. The table below lists interesting axioms and the criteria they characterize in the class of BD-frames.

C	$\vdash \mathrm{D}\,\varphi \rightarrow \neg\,\mathrm{B}\,\varphi$	**Mc**	$\forall w \in W,\ R^{\mathrm{B}}[w] \cap R^{\mathrm{D}}[w] \neq \emptyset$
BD	$\vdash \mathrm{B}\,\neg\varphi \rightarrow \mathrm{D}\,\varphi$	**Mbd**	$R^{\mathrm{D}} \subseteq R^{\mathrm{B}}$
DB	$\vdash \mathrm{D}\,\neg\varphi \rightarrow \mathrm{B}\,\varphi$	**Mdb**	$R^{\mathrm{B}} \subseteq R^{\mathrm{D}}$
Intro1	$\vdash \mathrm{D}\,\varphi \rightarrow \mathrm{B}\,\mathrm{D}\,\varphi$	**MI-1**	$wR^{\mathrm{B}}w' \wedge wR^{\mathrm{D}}w'' \Rightarrow w'R^{\mathrm{D}}w''$
Intro2	$\vdash \mathrm{B}\,\varphi \rightarrow \mathrm{D}\,\neg\,\mathrm{B}\,\varphi$	**MI-2**	$wR^{\mathrm{D}}w' \wedge wR^{\mathrm{B}}w'' \Rightarrow w'R^{\mathrm{B}}w''$

3 A comparative study

We have seen a plethora of frameworks that consider disbelief as a separate epistemic attitude in addition to belief. The natural question that arises here is: Are there any correspondences between this variety of frameworks? In this section we give a partial answer to this question based on the various studies that have been done over the years.

3.1 Belief-disbelief logics in a bi-logic framework

A detailed analysis of the various properties of the different logics presented in [3] in terms of the bi-logic framework can be found in [4]. We now consider a few of those properties to show the inter-dependence of the frameworks of [3].

1. If $\Gamma_D \vdash_2 \alpha$ and $\Gamma_D \vdash_2 \beta$ then $\Gamma_D \vdash_2 \alpha \vee \beta$. (disjunctiveness)
2. If $\Gamma_D \vdash_2 \alpha$ then $\neg\Gamma_D \vdash_1 \neg\alpha$. (role of negation-I)
3. If $\Gamma_B \vdash_1 \beta$ for all β then $\neg\Gamma_B \vdash_2 \beta$ for all β. (belief-disbelief consistency)
4. If $\Gamma_D \vdash_2 \alpha$ then $\Gamma_D \cup \{\neg\alpha\} \vdash_2 \beta$ for all β. (disbelief consistency)
5. If $\Gamma_D \vdash_1 \neg\alpha$ then $\Gamma_D \vdash_2 \alpha$. (role of negation-II)

These properties state interconnections between the dual consequence relations $\vdash_1$ and $\vdash_2$ and the negation in the bi-logic framework. The summary of the analysis is presented in the following table.

$\Vdash$	1	2	3	4	5
WBD	$\times$	$\surd$	$\times$	$\times$	$\times$
GBD	$\surd$	$\surd$	$\surd$	$\surd$	$\times$
BD	$\times$	$\times$	$\surd$	$\times$	$\surd$

Inter-dependences:

- **WBD** and **GBD** - From column 1 it is clear that **GBD** is not contained in **WBD**, and it can be proved that **WBD** is contained in **GBD**.
- **WBD** and **BD** - Independence follows from columns 2 and 3.
- **GBD** and **BD** - Independence follows from columns 4 and 5.

3.2 Embedding Gomolinska's logic in a modal framework

It has been shown in [6] that the logic of acceptance and rejection [1] corresponds to the modal framework of beliefs and disbeliefs proposed in [6]. A brief discussion is provided in the following.

It is a well-known result that Moore's autoepistemic logic corresponds to the nonmonotonic modal logic weak $S5$, in particular the $K45$-logic (see [19] for a detailed discussion). A similar correspondence has been shown between AEL2 [1] and a modal framework of belief and disbelief that is very close in spirit to what is proposed in [6]. Consider the following axioms:

$\mathbf{B_1}$	$B(\varphi \to \psi) \to (B\varphi \to B\psi)$	$\mathbf{D_1}$	$D(\varphi \wedge \psi) \to (D\neg\varphi \to D\psi)$
$\mathbf{B_2}$	$B\varphi \to BB\varphi$	$\mathbf{D_2}$	$D\varphi \to \neg DD\varphi$
$\mathbf{B_3}$	$\neg B\varphi \to B\neg B\varphi$	$\mathbf{D_3}$	$\neg D\varphi \to \neg D\neg D\varphi$

Let the underlying language be that of $\mathbf{L}_{BD}$. Let L_1 be the logic axiomatised by the propositional tautologies, B_1, B_2, and B_3, MP and B- necessitation rule, and let L_2 be the logic axiomatised by the propositional contradictions, D_1, D_2, and D_3, Rej and D-necessitation rule. The basic idea is to simulate the consequence closure and the groundedness properties of the stable expansions in AEL2. The notion of BD-dual extensions of a pair of accepted and rejected sets of formulas is defined in the following.

Definition 3.1. *(T_1, T_2) is said to be a BD-dual extension of (I_1, I_2) if,*
$T_1 = \{\psi \mid I_1 \cup \{\neg B\varphi : \varphi \notin T_1\} \cup \{D\varphi : \varphi \in T_2\} \cup \{\neg D\varphi : \varphi \notin T_2\} \vdash_{\mathsf{L}_1} \psi\}$;
$T_2 = \{\psi \mid I_2 \cup \{\neg B\varphi : \varphi \in T_1\} \cup \{B\varphi : \varphi \notin T_1\} \cup \{\neg D\varphi : \varphi \in T_2\} \vdash_{\mathsf{L}_2} \psi\}$.

The following proposition [6] brings out the equivalence of the frameworks.

Proposition 3.1. *(T_1, T_2) is a consistent stable* AEL2 *expansion of (I_1, I_2) iff (T_1, T_2) is a BD-dual extension of (I_1, I_2).*

3.3 Embedding different logics of belief and disbelief in a modal framework

A semantic correspondence (shown in [6]) is now provided between the logics developed in [3] and the modal frameworks developed in [6]. One should first note that in the modal language proposed in [6], an agent can have beliefs and disbeliefs not only about propositional ontic facts, but also about her own beliefs and disbeliefs.

For a semantic comparison, note how in [3], the agent's beliefs are given by M: the set of worlds (situations) she considers possible. She believes everything that is true in all the worlds (situations) in M, just like a standard modal operator. This coincides with the positive attitudes of the modal approach.

On the other hand, the agent's disbeliefs in [3] are not given by a single set of worlds, but by several of them: $\mathcal{N} = \{N_1, \ldots, N_k\}$ where k is finite. It is said that the agent disbelieves ϕ when ϕ is true in all worlds in N_i, for some $N_i \in \mathcal{N}$. More precisely, the agent's disbeliefs are not given by a single universal modality but by a combination of an existential and then a universal one.

Evidently, one can build up a Kripke model that encodes the information of a given **WBD** model. The idea behind the construction is the following. Consider a **WBD** model $\mathsf{M} = \langle M, \mathcal{N} \rangle$ in which $\mathcal{N}$ is finite, and denote by k its number of elements. An extension of the dual models (cf. Section 2.5) is built in which the domain consists of all the possible valuations for the given atomic propositions. Since the formulas are evaluated in pointed dual models, an evaluation world w is needed - any world in the dual model works. For representing beliefs, it is only needed to make R^B-accessible from w exactly those situations (valuations) the agent considers possible in M, that is, one should connect w by means of R^B, with all valuations in M. The representation of disbeliefs is different since the approach in [3] considers a set of sets of worlds. So first one fixes k worlds in the dual model, $\mathsf{w_1}, \ldots, \mathsf{w_k}$, each one of them standing for one of the sources that an agent has. Then, for every $i \in \{1, \ldots, k\}$, one makes R^D-accessible from $\mathsf{w_i}$ exactly the worlds the source i considers possible. What is left now is to indicate that the agent indeed has these k sources from the evaluation point w, and for this an auxiliary relation R^B_s is used that connects this w with every $\mathsf{w_i}$. Here follows a formal definition of the model construction.

Definition 3.2. *Let P be a set of atomic propositions and let $\mathsf{M} = \langle M, \mathcal{N} \rangle$ be a **WBD** model based on them, with $\mathcal{N} = \{N_1, \ldots, N_k\}$ (i.e, $\mathcal{N}$ is finite). Recall that V is the set of all propositional valuations over P, and denote by V_p the set of propositional valuations in V that make p true. The extended dual model $\mathcal{M}_{\mathsf{M}} = \langle W, R^B, R^B_s, R^D, V \rangle$ has as domain the set of all valuations for P, that is, $W := \mathsf{V}$. Now select arbitrary $k+1$ worlds $\mathsf{w}, \mathsf{w_1}, \ldots, \mathsf{w_k}$ in W.*

- *Define $R^B\mathsf{w}u$ iff $u \in M$. This way one indicates that, at w, the positive relation can reach exactly all the situations u the agent considers possible in the **WBD** model M.*

- *For each $i \in \{1, \ldots, k\}$, define $R^D \mathsf{w_i}' u$ iff $u \in N_i$. This way one indicates that, at each $\mathsf{w_i}$, the negative relation R^D can reach exactly all the situations u the negative source i considers possible in the **WBD** model M.*
- *Define $R^B_s \mathsf{w}\mathsf{w_i}$ for every $i \in \{1, \ldots, k\}$. This way one states that w can reach exactly all the $\mathsf{w_i}$ worlds (R^B_s is the* source *relation).*
- *For every atomic proposition p, define $V(p) := \mathsf{V}_p$.*

The properties of a **WBD** model that give rise to the logics **WBD**, **GBD**, **BD** and **BN** are directly encoded.

- A model $\mathsf{M} = \langle M, \mathcal{N} \rangle$ for the logic **GBD** has been defined as a **WBD** model in which $\mathcal{N}$ is a singleton. This produces a model $\mathcal{M}_\mathsf{M}$ in which $R^B_s[\mathsf{w}]$ is a singleton.
- A model $\mathsf{M} = \langle M, \mathcal{N} \rangle$ for the logic **BD** has been defined as a **WBD** model in which every $N \in \mathcal{N}$ is a subset of M. This produces a model $\mathcal{M}_\mathsf{M}$ in which $R^D[\mathsf{w_i}]$ is a subset of $R^B[\mathsf{w}]$ for every $i \in \{1, \ldots, k\}$.
- A model $\mathsf{M} = \langle M, \mathcal{N} \rangle$ for the logic **BN** has been defined as a model for the logic **BD** with the additional property that M is a subset of every $N \in \mathcal{N}$. This produces a model $\mathcal{M}_\mathsf{M}$ where, besides having $R^D[\mathsf{w_i}] \subseteq R^B[\mathsf{w}]$ for every $i \in \{1, \ldots, k\}$, we have $R^B[\mathsf{w}] \subseteq R^D[\mathsf{w_i}]$. This of course implies that $R^D[\mathsf{w_i}] = R^B[\mathsf{w}]$, for every $i \in \{1, \ldots, k\}$.

For this special dual model, a modality is used for each relation. The relations R^B, R^B_s are positive and R^D is negative, that is, for every world $w \in W$,

$$
\begin{array}{lll}
(\mathcal{M}_\mathsf{M}, w) \models B\,\varphi & \text{iff} & \text{for all } w' \text{ such that } R^B ww',\ (\mathcal{M}_\mathsf{M}, w') \models \varphi \\
(\mathcal{M}_\mathsf{M}, w) \models B_s\,\varphi & \text{iff} & \text{for all } w' \text{ such that } R^B_s ww',\ (\mathcal{M}_\mathsf{M}, w') \models \varphi \\
(\mathcal{M}_\mathsf{M}, w) \models D\,\varphi & \text{iff} & \text{for all } w' \text{ such that } R^D ww',\ (\mathcal{M}_\mathsf{M}, w') \models \neg\varphi
\end{array}
$$

The following proposition is immediate.

Proposition 3.2. *Let P be a set of atomic propositions and let $\mathsf{M} = \langle M, \mathcal{N} \rangle$ be a **WBD** model based on them, with $\mathcal{N}$ finite. For every propositional formula γ, it can be shown*

$$
\begin{array}{lll}
\mathsf{M} \models \gamma & \textit{iff} & (\mathcal{M}_\mathsf{M}, \mathsf{w}) \models B\,\gamma \\
\mathsf{M} \models \overline{\gamma} & \textit{iff} & (\mathcal{M}_\mathsf{M}, \mathsf{w}) \models \widehat{B_s}\,D\,\gamma
\end{array}
$$

The particular semantic requirements of the logics **GBD**, **BD** and **BN** correspond to the same semantic requirements as for extended dual models.

We have already seen that modeling the notion of disbelief in terms of possible worlds and accessibility relations, similar to the way one models beliefs can provide a uniform framework of belief and disbelief [6]. The comparative study that has been presented in this section shows that such a modal framework may capture all the intuitive properties of an agent's disbelief. Thus we can use the results of modal logic to investigate the finer inter-dependences between belief and disbelief.

4 Believing and disbelieving: removal of uncertainties

It is very natural for an agent to have uncertainties over certain propositions, which could be ideally represented as agents having both belief and disbelief or neither regarding certain facts. One can even think of having both belief and disbelief over certain facts as an inconsistent state of mind of the agent. We have already described a framework (see Section 2.3) which provides a fine-tuned procedure for removal of inconsistencies/uncertainties in an agent's mind by some choices that the agent makes in the course of her derivation process. In this section the preliminaries of a dynamic framework are introduced that can deal with such removal processes, in the spirit of dynamic epistemic logics [20]. We should mention here that, in this case also, we are giving an overview of the work done in [6].

The system proposed in Section 2.5 represents belief and disbelief by means of two modalities to construct formulas of the form $B\varphi$ and $D\varphi$. While the first one is true at a world w iff φ is true in all the worlds R^B– reachable from w, the second is true at a world w iff φ is *false* in all the worlds R^D– reachable from w. When the two relations are same, one gets the validity $D\varphi \leftrightarrow B\neg\varphi$ indicating that the agent disbelieves a proposition iff she believes in its negation. This actually says that, when $R^B = R^D$, disbelief collapses into classical negation, and therefore we get the classical normal modal logic system, K.

But from a more dynamic perspective, the case in which $R^B = R^D$, that is the K case, can be thought of not as a particular case of the static system, but as a possible result of some dynamic extension. In other words, the *ideal* system K in which disbeliefs coincide with classical negation, can be seen not as the state of an ideal static agent, but as the possible final state of a non-ideal but dynamic one who can perform actions that make the two relations the same. This section looks briefly at possible results of such actions.

There are various ways to generate a new relation from two others, and in this case they represent the different policies through which the agent 'merges' her beliefs and disbeliefs. For example, she can be drastic in two different ways: give up her negative attitude ($R := R^B$) or give up her positive attitude ($R := R^D$). More reasonable are the policies that actually combine the two relations, like $R := R^B \cup R^D$.

Definition 4.1 (Merging policies). *Let $\mathcal{M} = \langle W, R^B, R^D, V\rangle$ be a dual model. The relation R of a dual model $\langle W, R, V\rangle$ that results from the agent's merging of her beliefs and disbeliefs can be defined in several forms:*

- $R := R^B$ *(the* drastic positive *policy; new model denoted by $\mathcal{M}_B$).*
- $R := R^D$ *(the* drastic negative *policy; new model denoted by $\mathcal{M}_D$).*
- $R := R^B \cup R^D$ *(the* liberal combining *policy; new model denoted by $\mathcal{M}_\cup$).*
- $R := \alpha(R^B, R^D)$*, where $\alpha(R^B, R^D)$ is a PDL-expression [21] based on R^B and R^D (the* PDL *policy; new model denoted by $\mathcal{M}_\alpha$).*
- $R := R^B \cap R^D$ *(the* skeptic combining *policy; new model denoted by $\mathcal{M}_\cap$).*

– $R := R^B \setminus R^D$ *(new model denoted by* $\mathcal{M}_{\pm}$*).*
– $R := R^D \setminus R^B$ *(new model denoted by* $\mathcal{M}_{\mp}$*).*

For each policy $\circ$, a modality $[m_\circ]$ can be defined for building formulas of the form $\langle m_\circ \rangle \varphi$, read as *there is a way of merging attitudes with policy* $\circ$ *after which* φ *is the case*. Their semantic interpretation is given by:

$$(\mathcal{M}, w) \models \langle m_\circ \rangle \varphi \quad \text{iff} \quad (\mathcal{M}_\circ, w) \models \varphi$$

Now, for an axiom system, one can provide reduction axioms for each such policy. In each case, the relevant ones are those describing the way the new relations are created.

The *drastic positive* policy:	$\langle m_B \rangle \widehat{B}\varphi \leftrightarrow \widehat{B}\langle m_B \rangle \varphi$
	$\langle m_B \rangle \widehat{D}\varphi \leftrightarrow \widehat{B}\langle m_B \rangle \neg\varphi$
The *drastic negative* policy:	$\langle m_D \rangle \widehat{B}\varphi \leftrightarrow \widehat{D}\langle m_D \rangle \neg\varphi$
	$\langle m_D \rangle \widehat{D}\varphi \leftrightarrow \widehat{D}\langle m_D \rangle \varphi$
The *liberal combining* policy:	$\langle m_\cup \rangle \widehat{B}\varphi \leftrightarrow \widehat{B}\langle m_\cup \rangle \varphi \vee \widehat{D}\langle m_\cup \rangle \neg\varphi$
	$\langle m_\cup \rangle \widehat{D}\varphi \leftrightarrow \widehat{B}\langle m_\cup \rangle \neg\varphi \vee \widehat{D}\langle m_\cup \rangle \varphi$

For $*$-free *PDL* policies, reduction axioms for each particular $\alpha(R^B, R^D)$ can be obtained by following the technique introduced in [22]. To get sound and complete reduction axioms for the policies involving $\cap$ and $\setminus$, it may be needed to extend the language with nominals.

But we do not only need to look at actions that create a single relation in one shot. We can also look at procedures in which the single relation is a long-term result of small operations that merge the two of them in a step-wise form. This will be more in the line of the procedural framework described in Section 2.3. Consider the following issues, i) analyzing the merging policies in more detail, and ii) constructing a dynamic framework that can describe the smaller steps taken in achieving the ultimate goal, that is of removing inconsistency. These are some natural research questions that can be investigated further.

5 Concluding remarks

In this contribution to the festschrift volume for Mihir Chakraborty, we have attempted to provide a fresh look at the different frameworks of disbelief which has been a topic of his interest over the years. As it is evident from the discussions above, there are various loose ends to tie up, specially with respect to the study on the inter-dependence between belief and disbelief. Another interesting area of future research constitutes the study of the dynamics of removing uncertainties or inconsistencies present in an agent's epistemic state.

Acknowledgements: We thank Rineke Verbrugge for patiently reading a preliminary version of this article. Her comments helped us to give this article its present form.

References

1. Gomolinska, A.: On the logic of acceptance and rejection. Studia Logica **60**(2) (1998) 233–251
2. Ghose, A., Goebel, R.: Belief states as default theories: Studies in non-prioritized belief change. In: Proceedings of the European Conference on Artificial Intelligence. (1998) 8–12
3. Chopra, S., Heidema, J., Meyer, T.A.: Some logics of belief and disbelief. In Gedeon, T.D., Fung, L.C.C., eds.: Australian Conference on Artificial Intelligence. Volume 2903 of Lecture Notes in Computer Science., Springer (2003) 364–376
4. Ghosh, S.: Nonmonotonic Reasoning and Belief Revision – Two Aspects of Common Sense Reasoning: A Logic-Algebraic Study. PhD thesis, University of Calcutta, Kolkata, India (2005)
5. de Jongh, D., Ghosh, S.: Comparing strengths of beliefs explicitly. In Baldoni, M., et al., eds.: MALLOW. Volume 494 of CEUR Workshop Proceedings., CEUR-WS.org (2009)
6. Ghosh, S., Velázquez-Quesada, F.: A qualitative approach to uncertainty. In Banerjee, M., Seth, A., eds.: Proceedings of the 4th Indian Conference on Logic and its Applications (ICLA 2011). Volume 6521, LNAI. (2011) 90–104
7. Chakraborty, M.: Ganiter dharapat o galposhalpo. Nandimukh Samsad (2001)
8. Kyburg, H.: Probability and the Logic of Rational Belief. Wesleyan University Press, Middletown, CT (1961)
9. Arló-Costa, H.: Non-adjunctive inference and classical modalities. The Journal of Philosophical Logic **34** (2005) 581–605
10. Lewis, D.: Elusive knowledge. Australasian Journal of Philosophy **74**(4) (1996) 549–567
11. Gomolinska, A., Pearce, D.: Disbelief change. Electronic essays on the occasion of the fiftieth birthday of Peter Gardenfors (2001)
12. Tversky, A., Kahneman, D.: Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty **5**(4) (1992) 297–323
13. Caminada, M.W.A., Gabbay, D.M.: A logical account of formal argumentation. Studia Logica **93**(2-3) (2009) 109–145
14. Dubois, D., Prade, H., eds.: Special issue on *Bipolarity.* (2008) International Journal of Intelligent Systems, **23**(8, 9, 10).
15. Dubois, D., Prade, H.: Foreword. International Journal of Intelligent Systems **23**(8) (2008) 863–865
16. Moore, R.C.: Semantical considerations on nonmonotonic logic. Artificial Intelligence **25**(1) (1985) 75–94
17. Alchourrón, C.E., Gärdenfors, P., Makinson, D.: On the logic of theory change: Partial meet contraction and revision functions. The Journal of Symbolic Logic **50**(2) (1985) 510–530
18. Halpern, J.Y., Moses, Y.: A guide to completeness and complexity for modal logics of knowledge and belief. Artificial Intelligence **54**(2) (1992) 319–379
19. Schwarz, G.: Autoepistemic modal logics. In: Proceedings of the 3rd Conference on Theoretical Aspects of Rationality and Knowledge (TARK). (1990)
20. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Volume 337 of Synthese Library. Springer (2007)
21. Harel, D., Kozen, D., Tiuryn, J.: Dynamic Logic. MIT Press, Cambridge (2000)
22. van Benthem, J., Liu, F.: Dynamic logic of preference upgrade. Journal of Applied Non-Classical Logics **17**(2) (2007) 157–182

Sujata Ghosh
Department of Artificial Intelligence,
University of Groningen, Netherlands.
sujata@ai.rug.nl

Topological systems, Topology and Frame: in fuzzy context

Purbita Jana

Topology via Logic [2] is an inspiring work. This book gives some idea about the interconnection between Topological systems, Topology and Frame. We are investigating a similar interconnection with regard to Fuzzy Topological System, Fuzzy Topology and Frame.

Interconnection between Boolean Systems, Topology and Boolean Algebra was investigated in my project work at the Institute for Logic, Language and Computation, Amsterdam. That work gave us Stone duality as a result. We now want to see whether it can be extended to the fuzzy case as well.

The idea present in the work in "Fuzzy Topology and Łukasiewicz Logics from the Viewpoint of Duality Theory" [1], where "fuzzy topological duality for the algebras of Łukasiewicz n-valued truth constant" is established, is somehow similar to ours.

We have already worked out n-Fuzzy Boolean System. Further, we have investigated interconnection between n-Fuzzy Boolean System and Łukasiewicz n-valued logic with truth constants. Interconnection between n-Fuzzy Boolean System and n-Fuzzy Boolean Space has also been investigated.

References

1. Maruyama, Y.: Fuzzy Topology and Łukasiewicz Logics from the Viewpoint of Duality Theory. Studia Logica. pp 245-269, (2010).
2. Vickers, S.: Topology via Logic. (Vol. 5). Cambridge: Cambridge Tracts in Theoretical Computer Science University Press, (1989).

Purbita Jana
Supreme Knowledge Foundation Group of Institutions,
Hooghly, West Bengal
purbita.jana@gmail.com

Generalizations and Logics of Rough Set Theory*

Md. Aquil Khan

1 Introduction

Rough set theory, introduced by Pawlak, is based on the concept of *approximation space* [35] which is defined as a tuple (W, R), where R is an equivalence relation on the set W. Any concept represented as a subset (say) X of the partitioned domain W, is then approximated from 'within' and 'outside', by its *lower* and *upper approximations* given as $\underline{X}_R := \{x : [x] \subseteq X\}$ and $\overline{X}_R := \{x : [x] \cap X \neq \emptyset\}$ respectively. Here, $[x]$ denotes the equivalence class of $x \in W$. Given an approximation space (W, R) with a finite W, the *rough membership function* $f : W \times 2^W \to [0, 1]$ is defined as $f(x, X) = \dfrac{|[x] \cap X|}{|[x]|}$, $x \in W,\ X \subseteq W$. Note that $f(x, X) = 1$ if and only if $x \in \underline{X}_R$, and $f(x, X) = 0$ if and only if $x \notin \overline{X}_R$.

A practical source of a Pawlak approximation space is an *information system* [35], formally defined as a tuple $\mathcal{S} := (W, A, \{Val_a\}_{a\in A}, f)$, consisting of a non-empty set W of objects, a non-empty set A of attributes, a non-empty set Val_a of attribute values for each $a \in A$, and $f : W \times A \to \bigcup_{a\in A} Val_a$ such that $f(x, a) \in Val_a$. Any information system $\mathcal{S} := (W, A, \{Val_a\}_{a\in A}, f)$ and $B \subseteq A$ would induce an 'indiscernibility' relation $Ind_{\mathcal{S}}(B)$ on W:

$x\ Ind_{\mathcal{S}}(B)\ y$ if and only if $f(x, a) = f(y, a)$ for all $a \in B$.

Thus, given an information system $\mathcal{S}$, and a set B of attributes, we obtain an approximation space $(U, Ind_{\mathcal{S}}(B))$.
Note that the above notion of information system can be generalized by considering the function f to be a mapping from $W \times A$ to $2^{\bigcup_{a\in A} Val_a}$. Such an information system is called *non-deterministic information system*. A non-deterministic information system induces the following relations apart from the indiscernibility.

(Similarity) $x\ Sim_{\mathcal{S}}(B)\ y$ if and only if $f(x, a) \cap f(y, a) \neq \emptyset$ for all $a \in B$.
(Inclusion) $x\ In_{\mathcal{S}}(B)\ y$ if and only if $f(x, a) \subseteq f(y, a)$ for all $a \in B$.

Since the inception of rough set theory, a large number of high quality papers on various aspects of rough sets and their applications have been published. In this article, our aim is to present a survey of the works related to two important directions of research on rough set theory, viz. generalizations of Pawlak's rough set approach, and logical aspects of rough set theory.

* I dedicate this article to our beloved teacher M. K. Chakraborty

2 Generalized rough set models

Over the years it has been observed that Pawlak's simple idea of rough set theory needs extensions to make it applicable in different practical situations. As a consequence of it, one can find several generalizations of rough set theory in literature. In this section, we present some of these generalizations.

2.1 Approximation space with relation other than equivalence

Most natural generalization is the one where the relation R is not necessarily an equivalence. So different types of R gives different useful generalizations of the Pawlak rough set model. For instance, in [41, 25], tolerance approximation space is considered which is a tuple (U, R) with tolerance, i.e with a reflexive and symmetric relation R. The notion of lower and upper approximations in these generalized approximation spaces are defined in standard way as follows. Let $R(x) := \{y \in U : (x, y) \in R\}$. Then for $X \subseteq U$, $\underline{X}_R := \{x \in U : R(x) \subseteq X\}$ and $\overline{X}_R := \{x \in U : R(x) \cap X \neq \emptyset\}$.

2.2 Multiple relation approximation space

Another natural generalization of approximation space is one where we consider a number of relations instead of just one. In [35], a relational system $K := (U, \mathbf{R})$ is considered, where $\mathbf{R}$ is a family of equivalence relations. Moreover, for any $P \subseteq \mathbf{R}$ and $P \neq \emptyset$, the equivalence relation $IND(P)$, which is the intersection of all equivalence relations belonging to P, is considered. One can view $\mathbf{R}$ as the collection of indiscernibility relations corresponding to individual attributes. So, for $P \subseteq \mathbf{R}$, $IND(P)$ is the indiscernibility relation corresponding to the set of attributes consisting of precisely those attributes, the corresponding indiscernibility relation of which are in P. This intuition is more explicitly represented by *information structure* proposed by Orłowska [30]. It is a structure of the form $(U, \{R_B\}_{B\subseteq\mathcal{A}})$, where $\mathcal{A}$ is a non-empty set of parameters or attributes and for each $B \subseteq \mathcal{A}$, R_B is an equivalence relation on U satisfying,

$$R_\emptyset = U \times U \tag{1}$$

$$R_{B\cup C} = R_B \cap R_C. \tag{2}$$

Condition 1 signifies that we can distinguish objects only using the information about the objects regarding the attributes. We note that given an information system $\mathcal{S} := (W, \mathcal{A}, \{Val_a\}_{a\in A}, f)$, the structure $(W, \{Ind_{\mathcal{S}}(B)\}_{B\subseteq\mathcal{A}})$ is an information structure. For every information structure $(W, \{R_B\}_{B\subseteq\mathcal{A}})$, can we determine an information system $\mathcal{S} := (W, \mathcal{A}, \{Val_a\}_{a\in A}, f)$ such that $Ind_{\mathcal{S}}(B) = R_B$ for all $B \subseteq \mathcal{A}$? The answer is yes, *provided* $\mathcal{A}$ *is finite*. This is due to the fact that an information structure may not have the property $R_B = \bigcap_{b\in B} R_{\{b\}},\ B \subseteq \mathcal{A}$ as shown in Example 1 below, but we always have $Ind_{\mathcal{S}}(B) = \bigcap_{b\in B} Ind_{\mathcal{S}}(\{b\})$.

Example 1. Consider $\mathfrak{F} := (U := \{x, y\}, \{R_B\}_{B \subseteq \mathcal{A}})$ where $R_B := U \times U$ for any finite subset B of $\mathcal{A}$, while $R_B := Id_U$ for infinite B. Note that for any infinite B, we have $R_B \neq \bigcap_{b \in B} R_{\{b\}}$.

Given the collection of indiscernibility relations corresponding to different sets of attributes, one can define different operations on this collection to obtain new properties. For instance, in [11], the intersection and transitive closure of union of indiscernibility relations is considered. A structure of the form $(U, \{R_a\}_{a \in REL})$ is defined, where REL is a set of relational expressions built inductively using a set $\mathcal{R}$ of relational variables and two binary operations $\cap$ and $\uplus$ such that

- for each $a \in REL$, R_a is an equivalence relation on U,
- for $a, b \in REL$, $R_{a \cap b} = R_a \cap R_b$ and $R_{a \uplus b} = (R_a \cup R_b)^+$ (transitive closure of $R_a \cup R_b$).

Let us call the above structure, a *data analysis logic structure*, or in brief, a *DAL-structure*. A variant of DAL-structure is considered in [12], where $R_{a \uplus b}$ is taken as $R_a \cup R_b$, keeping all the other conditions same as in DAL-structure. Let us call it a *$DALLA$-structure*. So every $DALLA$-structure is a DAL-structure, but not conversely.

Collection of equivalence relations over the same domain is also considered in [39], but the motivation is different from the structures discussed earlier. Multi-agent scenario is considered where each agent has its own knowledge base represented by an equivalence relation. Thus in [39], a structure of the form $(U, \{R_t\}_{t \in \mathcal{T}})$ is considered, where $\mathcal{T}$ is a set of terms built using a set T of individual agents and two binary operations $\wedge$ and $\vee$ such that

- for each $a \in \mathcal{T}$, R_a is an equivalence relation on U,
- for $a, b \in \mathcal{T}$,
  $U/R_{a \vee b} := \{[x]_{R_a} \cap [y]_{R_b} : [x]_{R_a} \cap [y]_{R_b} \neq \emptyset\}$ and
  $U/R_{a \wedge b} := \{[x]_{R_a} \cup [y]_{R_b} : [x]_{R_a} \cap [y]_{R_b} \neq \emptyset\}$.

If R_a and R_b represents the knowledge base of the agents a and b respectively, then $R_{a \vee b}$ and $R_{a \wedge b}$ are respectively called the *strong distributed knowledge base* and *weak distributed knowledge base* of the group $\{a, b\}$ of agents. Note that $R_{a \vee b} = R_a \cap R_b$ and hence notion of strong distributed knowledge base can be identified with the notion of distributed knowledge in epistemic logic [8].

Multiple agent scenario is also considered in [23], although a more general term 'source' is used. The notion of *multiple-source approximation system with distributed knowledge base* ($MSAS^D$) is used to represent such a situation. It is given by a tuple $\mathfrak{F} := (U, \{R_P\}_{P \subseteq N})$, where U is a non-empty set, N an initial segment of the set $\mathbb{N}$ of positive integers and for each $P \subseteq N$, R_P is a binary relation on U satisfying the following:

(M1) R_P is an equivalence relation;
(M2) $R_P = \bigcap_{i \in P} R_i$, for each $P \subseteq N$.

For $i \in N$, $R_{\{i\}}$ denotes the knowledge base of the i^{th} source of the system. It is observed that, either of the following approaches can be taken to obtain notions of approximations in the multiple source situation. One may define upper and lower approximations with respect to each source, and then combine them (via conjunction or disjunction). On the other hand, one may also define a new relation using the (equivalence) relations attached to each source and then construct the upper and lower approximations with respect to this new relation. Note that the approximations with respect to strong/weak distributed knowledge base considered in [39] are instances of this approach. Moreover, the approximations defined in [24] are also based on this idea. We shall discuss it in Sect. 2.5. The notions of strong/weak lower and upper approximations proposed in [23], on the other hand, are based on the first approach. Let us consider a $MSAS^D$ $\mathfrak{F} := (U, \{R_P\}_{P \subseteq N})$. Then, corresponding to each $P(\neq \emptyset) \subseteq N$, the notions of strong lower approximation $\underline{X}_{s(P)}$, weak lower approximation $\underline{X}_{w(P)}$, strong upper approximation $\overline{X}_{s(P)}$, and weak upper approximation $\overline{X}_{w(P)}$ are defined as follows.

$$\underline{X}_{s(P)_{\mathfrak{F}}} := \bigcap_{i \in P} \underline{X}_{R_i}; \;\; \underline{X}_{w(P)_{\mathfrak{F}}} := \bigcup_{i \in P} \underline{X}_{R_i}.$$
$$\overline{X}_{s(P)_{\mathfrak{F}}} := \bigcap_{i \in P} \overline{X}_{R_i}; \;\; \overline{X}_{w(P)_{\mathfrak{F}}} := \bigcup_{i \in P} \overline{X}_{R_i}.$$

Observe that in the special case when $|N| = 1$, the weak/strong lower and upper approximations are just the standard Pawlak's lower and upper approximations respectively.

In the above considered structures, only a collection of equivalence relations over the same domain is considered. This can be further generalized by considering a collection of relations other than equivalence, or even a collection of different types of relations over the same domain. For instance, in [33], a structure of the form $(U, \{R_i\}_{i \in I})$, called dynamic space, is considered, where $\{R_i\}_{i \in I}$ is a family of binary relations. On the other hand, in [42], *NIL-structure* (non-deterministic information logic structure) is defined which is a structure of the form (U, S, R), where R and S are binary relations on U satisfying the following.

1. $(x, x) \in R \cap S$.
2. if $(x, y) \in R$ and $(y, z) \in R$ then $(x, z) \in R$.
3. If $(x, y) \in S$, then $(y, x) \in S$.
4. If $(x, y) \in S$, $(x, u) \in R$, $(y, v) \in R$, then $(u, v) \in S$.

Note that given a non-deterministic information system $\mathcal{K} := (U, \mathcal{A}, \{Val_a\}_{a \in \mathcal{A}}, f)$, $(U, Sim_{\mathcal{K}}(\mathcal{A}), In_{\mathcal{K}}(\mathcal{A}))$ is a NIL-structure, called *standard NIL-structure*. Moreover, in [42], the following proposition is proved.

Proposition 1. *Every NIL-structure is standard, that is, given a NIL-structure (U, S, R), there exists a non-deterministic information system $\mathcal{K} := (U, \mathcal{A}, \{Val_a\}_{a \in \mathcal{A}}, f)$ such that $Sim_{\mathcal{K}}(\mathcal{A}) = S$ and $In_{\mathcal{K}}(\mathcal{A}) = R$.*

2.3 Rough set models based on covering and neighborhood system

Pawlak approximation space can also be generalized by considering a *covering* of the domain instead of partition. By a covering of a domain U, we mean a collection of subsets $\{C_i\}$ of U such that $\bigcup C_i = U$. Thus, every partition is also a covering. Given a covering $\mathcal{C} := \{C_i\}$ of U, the following four types of lower and upper approximations are defined in [36]. Let $N_x^{\mathcal{C}} := \bigcup\{C_i : x \in C_i\}$ and $P_x^{\mathcal{C}} := \{y \in U : \forall C_i(x \in C_i \Leftrightarrow y \in C_i)\}$.

1. $\underline{\mathcal{C}}_1(X) := \{x : N_x^{\mathcal{C}} \subseteq X\}$.
   $\overline{\mathcal{C}}^1(X) := \bigcup\{C_i : C_i \cap X \neq \emptyset\}$.
2. $\underline{\mathcal{C}}_2(X) := \bigcup\{N_x^{\mathcal{C}} : N_x^{\mathcal{C}} \subseteq X\}$.
   $\overline{\mathcal{C}}^1(X) := \{z : \forall x(z \in N_x^{\mathcal{C}} \Rightarrow N_x^{\mathcal{C}} \cap X \neq \emptyset)\}$.
3. $\underline{\mathcal{C}}_3(X) := \bigcup\{C_i : C_i \subseteq X\}$.
   $\overline{\mathcal{C}}^3(X) := \{y : \forall C_i(y \in C_i \Rightarrow C_i \cap X \neq \emptyset)\}$.
4. $\underline{\mathcal{C}}_4(X) := \bigcup\{P_x^{\mathcal{C}} : P_x^{\mathcal{C}} \subseteq X\}$.
   $\overline{\mathcal{C}}^4(X) := \bigcup\{P_x^{\mathcal{C}} : P_x^{\mathcal{C}} \cap X \neq \emptyset\}$.

Note that all of the above four notions of lower and upper approximations coincide with the Pawlak's standard notion of lower and upper approximations when the covering $\{C_i\}$ is actually a partition of U, i.e. when it satisfies the additional property $C_i \cap C_j \neq \emptyset \Rightarrow C_i = C_j$.

The above notion of approximation space based on covering is further generalized in [26]. A *neighborhood system* is considered in [26], defined as a tuple (U, N), where $N : U \to 2^{2^U}$ satisfies the following:

1. $\emptyset \notin N(x)$ for all $x \in U$,
2. for all $x \in U$ and $X \subseteq Y \subseteq U$, if $X \in N(x)$, then $Y \in N(x)$.

The lower and upper approximation of a set X, denoted by $\underline{N}(X)$ and $\overline{N}(X)$, is then defined as follows:

$$\underline{N}(X) := \{x \in U : \exists Y \in N(x), Y \subseteq X\}$$
$$\overline{N}(X) := \{x \in U : \forall Y \in N(x), Y \cap X \neq \emptyset\}.$$

Note that given a generalized approximation space (U, R), R being any binary relation, one can define the neighborhood system (U, N), where $N(x) := \{X \subseteq U : R(x) \subseteq X\}$. Moreover, we obtain $\underline{N}(X) = \underline{X}_R$ and $\overline{N}(X) = \overline{X}_R$.

Let $\mathcal{C} := \{C_i\}$ be a covering of a domain U. Let us consider the neighborhood system $(U, N^{\mathcal{C}})$, where $N^{\mathcal{C}}(x) := \{C_i : x \in C_i\}$. Then we have the following proposition relating the covering and neighborhood based rough set models.

Proposition 2. *1. Neighborhood system $(U, N^{\mathcal{C}})$ satisfies the following.*
   (a) $x \in X \in N^{\mathcal{C}}(y) \Rightarrow X \in N^{\mathcal{C}}(x)$;
   (b) $X \in N^{\mathcal{C}}(x) \Rightarrow x \in X$;
   (c) $N^{\mathcal{C}}(x) \neq \emptyset$ for all $x \in U$.
2. $\underline{N^{\mathcal{C}}}(X) = \underline{\mathcal{C}}_3(X)$ and $\overline{N^{\mathcal{C}}}(X) = \overline{\mathcal{C}}_3(X)$, where $\underline{\mathcal{C}}_3(X)$ and $\overline{\mathcal{C}}_3(X)$ are defined above.
3. Given a neighborhood system (U, N) satisfying (1a)-(1c), there exists a covering $\mathcal{C}$ of U such that $N^{\mathcal{C}} = N$.

2.4 Rough set models based on membership functions

In [45], variable precision rough set model (VPRS-model) is proposed which is based on the Pawlak approximation space (with finite domain) together with a generalized notion of lower and upper approximation. In fact, the notion of lower approximation is generalized by replacing the inclusion relation with a *majority inclusion relation* in the original definition of Pawlak lower approximation of a set. For $\beta \in [0, \frac{1}{2})$, majority inclusion relation $\subseteq^\beta$ is defined as follows:

$$X \subseteq^\beta Y \text{ if and only if } C(X,Y) \leq \beta,$$

where $C(X,Y) = 1 - \frac{|X \cap Y|}{|X|}$, if $|X| > 0$; otherwise, $C(X,Y) = 0$.
Note that $C([x]_R, X) = 1 - f([x]_R, X)$, f being the rough membership function. Now, using $\subseteq^\beta$, the β-lower approximation $\underline{R}_\beta X$ and β-upper approximation $\overline{R}_\beta X$ of a set $X \subseteq U$ is defined as follows:

$$\underline{R}_\beta X := \{x \in U : [x]_R \subseteq^\beta X\} = \{x \in U : f([x]_R, X) \geq 1 - \beta\},$$
$$\overline{R}_\beta X := \{x \in U : C([x]_R, X) < 1 - \beta\} = \{x \in U : f([x]_R, X) > \beta\}.$$

Observe that $\underline{R}_0 X = \underline{X}_R$ and $\overline{R}_0 X = \overline{X}_R$.
Note that the definitions of the membership function, β-lower and upper approximations are well defined when the domain U is finite. But these may not be well defined when U is infinite. This problem is dealt with by introducing *probabilistic approximation space* defined as the tuple (U, R, Pr), where (U, R) is a Pawlak approximation space and Pr is a probability distribution on U. In that case, one modify the rough membership function as

$$f(x,X) := \frac{Pr([x]_R \cap X)}{Pr([x]_R)}.$$

When U is finite and Pr is a uniform distribution, we just obtain the rough membership function. Now, by replacing the rough membership function with the above modified membership function in the definition of β-lower and upper approximations, we obtain the corresponding notions in the probabilistic approximation space (U, R, Pr).

The rough membership function is an absolute rough membership function in the sense that it takes into account the objects from the set $[x]_R$ only, and not the objects outside $[x]_R$. The relative rough membership functions are considered in [13] which depends on the objects from the set $[x]_R$ as well as the objects outside it. Different types of relative rough membership functions are defined in the following:

$d(x,X) := \frac{|[x]_R \cap X|}{|[x]_R|} - \frac{|X|}{|U|}$
$r(x,X) := log\left[\frac{|[x]_R \cap X|}{|[x]_R|} / \frac{|X|}{|U|}\right]$
$l(x,X) := log\left[\frac{|[x]_R \cap X|}{|X|} / \frac{|[x]_R \cap (U \setminus X)|}{|U \setminus X|}\right]$

$f(x,X) := \left[\frac{|[x]_R \cap X|}{|X|} - \frac{|[x]_R \cap (U \setminus X)|}{|U \setminus X|}\right] / \left[\frac{|[x]_R \cap X|}{|X|} + \frac{|[x]_R \cap (U \setminus X)|}{|U \setminus X|}\right]$
$s(x,X) := \frac{|[x]_R \cap X|}{|[x]_R|} - \frac{|(U \setminus [x]_R) \cap X|}{|U \setminus [x]_R|}$
$b(x,X) := \frac{|[x]_R \cap X|}{|U|} - \frac{|[x]_R|}{|U|}\frac{|X|}{|U|}.$

Using the relative membership function $C(x,X)$, the following parametrized lower and upper approximations of a set X is defined. Let t,q be two real numbers such that $0 \leq q \leq t \leq 1$ and $\alpha, \beta,\ \alpha \geq \beta$ be two real values in the range of variation of $C(x,X)$. Then

$$\underline{R}_{t,\alpha}(X) = \{x \in U : \frac{|[x]_R \cap X|}{|[x]_R|} \geq t \text{ and } C(x,X) \geq \alpha\},$$
$$\overline{R}_{q,\beta}(X) = \{x \in U : \frac{|[x]_R \cap X|}{|[x]_R|} > q \text{ and } C(x,X) > \beta\}.$$

Note that Pawlak rough set model is obtained when $t = 1, q = 0, \alpha = \beta = \min\{C(x,X) : x \in U \text{ and } X \subseteq U\}$. Moreover, VPRS model is obtained when $0 \leq q \leq t \leq 1, \alpha = \beta = \min\{C(x,X) : x \in U \text{ and } X \subseteq U\}$.

2.5 Fuzzy approximation space

In [7], the Pawlak approximations space is generalized by bringing the notion of fuzzy sets. A fuzzy approximation space is defined as a tuple (U,R), where R is now a fuzzy binary relation on U, i.e. $R : U \times U \to [0,1]$. The fuzzy lower and upper approximation $\underline{R}F, \overline{R}F : U \to [0,1]$ of a crisp or fuzzy subset F of U is defined as follows:

$$\underline{R}F(u) := \inf_{v \in U} R(u,v) \to_* F(v),$$

$$\overline{R}F(u) := \sup_{v \in U} R(u,v) * F(v),$$

where $* : [0,1] \times [0,1] \to [0,1]$ is a $t - norm$ and $\to_*$: $[0,1] \times [0,1] \to [0,1]$ is the S-implication with respect to $*$ defined by $a \to_* b := 1 - (a * (1 - b))$. In particular, when F is a crisp subset of U, the above two equations reduce to

$$\underline{R}F(u) := \inf_{v \notin F} 1 - R(u,v),$$

$$\overline{R}F(u) := \sup_{v \in F} R(u,v).$$

In order to handle multiple agent situation, [24] considered a special type of fuzzy approximation space, called $[0,1] \cap \mathbb{Q}$-*tolerance space*. It is a fuzzy approximation space $\mathfrak{G} := (U,R)$ satisfying the following additional conditions: for all $x, y \in U$,

- $R : U \times U \to [0,1] \cap \mathbb{Q}$.
- (reflexivity) $R(x,x) = 1$.
- (symmetry) $R(x,y) = R(y,x)$.

The following notions of apprpoximations based on $[0,1] \cap \mathbb{Q}$-tolerance spaces are proposed in [24]. For a $[0,1] \cap \mathbb{Q}$-fuzzy approximation space $\mathfrak{G} := (U, R)$, $x \in U$ and a *threshold* $\lambda \in (0,1] \cap \mathbb{Q}$, we will write $x \uparrow^{\lambda}$ to denote the set $\{y \in U : R(x,y) \geq \lambda\}$. The lower approximation $L_{\mathfrak{G}}$ and upper approximation $U_{\mathfrak{G}}$ of X of degree $\lambda \in (0,1] \cap \mathbb{Q}$ are then defined as follows:

$$L_{\mathfrak{G}}(X, \lambda) := \{x \in U : x \uparrow^{\lambda} \subseteq X\}, \quad U_{\mathfrak{G}}(X, \lambda) := \{x \in U : x \uparrow^{\lambda} \cap X \neq \emptyset\}.$$

We note that in a $[0,1] \cap \mathbb{Q}$-fuzzy approximation space $\mathfrak{G} := (U, R)$ with finite $R[U \times U]$, we obtain the following relationship between the notions of approximations defined in [7] and [24]:

$$x \in U_{\mathfrak{G}}(X, \lambda) \text{ if and only if } \overline{R}X(x) \geq \lambda.$$

3 Logics for indiscernibility relation

In this section, we will survey some of the logics defined for rough set theory. Some parts of this section has appeared in [4], where logics for rough set theory were surveyed. But here we restrict our study to the logics for the indiscernibility relations only. Moreover, it also includes some work done after the publication of the article [4].

3.1 Normal Modal Systems

The modal nature of the lower and upper approximations of rough sets was evident from the start. Hence, it is of no surprise that normal modal systems were focussed upon, during investigations on logics for rough sets. In particular, in case of Pawlak rough sets, the two approximations considered as operators clearly obey all the $S5$ laws. The formal connection between the syntax of $S5$ and its semantics in terms of rough sets is given as follows [3].

According to the Kripke semantics for $S5$, a well-formed formula (wff) α is interpreted by a function v as a subset in a non-empty domain U, the subset representing the extension of the wff – i.e. the collection of objects/worlds where the wff holds. Moreover, in a $S5$-model $\mathcal{M} := (U, R, v)$ (say), the accessibility relation R is an equivalence on U. Further, if $\Box$, $\Diamond$ denote the necessity and possibility operators respectively then for any wff α, $v(\Box\alpha) = \underline{v(\alpha)}_R$ and $v(\Diamond\alpha) = \overline{v(\alpha)}_R$.

A wff α is *true* in $\mathcal{M}$, if $v(\alpha) = U$. Now it can easily be seen that all the $S5$ theorems involving $\Box$ and $\Diamond$ translate into valid properties of lower and upper approximations.

Taking a cue from this connection, a multi-modal logic is defined in [31]. The language of the logic contains a set $CONREL$ of constants representing indiscernibility relations. Using the standard Boolean connectives, the set of all wffs is defined following the scheme: $p \in PV \mid \neg\alpha \mid \alpha \wedge \beta \mid [R]\alpha$, where PV is the set of propositional variables and $R \in CONREL$.

The semantics is based on a structure of the form $\mathfrak{F} := (U, \{R_i\}_{i\in I})$, where $\{R_i\}_{i\in I}$ is a family of equivalence relations over U. The satisfiability relation is defined using the meaning functions $m : CONREL \to \{R_i\}_{i\in I}$ and $v : PV \to 2^U$ in a standard way. For instance, for $\mathfrak{M} := (\mathfrak{F}, m, v)$ and $w \in U$,

$$\mathfrak{M}, w \models [R]\alpha \text{ if and only if for all } w' \text{ such that } (w, w') \in m(R),\ \mathfrak{M}, w' \models \alpha.$$

The collection $\{R_i\}_{i\in I}$ of equivalence relations, as mentioned in [31], is intended to represent the family of indiscernibility relations corresponding to a family of information systems over the same domain.

A sound and complete deductive system for the logic is also presented consisting of the following axioms.

1. $[R](\alpha \to \beta) \to ([R]\alpha \to [R]\beta)$.
2. $[R]\alpha \to \alpha$.
3. $\alpha \to [R]\langle R\rangle\alpha$, where $\langle R\rangle\alpha := \neg[R]\neg\alpha$.
4. $[R]\alpha \to [R][R]\alpha$.

Observe that this axiomatic system is the same as the axiomatic system of epistemic logic $S5_n$ without the common knowledge and distributed knowledge operators [8]. However, there is a difference in the language of the two logics. $S5_n$ has n modal operators representing the epistemic state of n agents. On the other hand, nothing is said about the cardinality of $CONREL$. On the side of semantics, in case of $S5_n$, we have structures with exactly n equivalence relations. But in the case of the logic defined in [31], we do not have any such restriction.

3.2 Logic *DAL* and its variant

As mentioned in Sect. 2.2, one can obtain different properties by defining operations on the set of indiscernibility relations, and we obtained a structure such as DAL-structure (cf. Sect. 2.2). A logic DAL (data analysis logic) for DAL-structure is proposed in [11]. The language of DAL, as in the case of the logic defined in [31], contains a set $\mathcal{R}$ of relation variables representing indiscernibility relations. Moreover, there are binary operations $\cap, \uplus$, and a collection REL of *relational expressions* that are built inductively out of the members of $\mathcal{R}$ with these operations. Thus the set of all DAL wffs is defined following the scheme: $p \in PV \mid \neg\alpha \mid \alpha\wedge\beta \mid [a]\alpha$, where $a \in REL$. For a DAL-structure $(U, \{R_a\}_{a\in REL})$, the satisfiability relation is defined in a similar way as in case of the logic in [31], using a meaning function $m : REL \to \{R_a\}_{a\in REL}$ such that $m(a) = R_a$.

As mentioned in [6], DAL is the paradigm logic for reasoning about indiscernibility relation, but unfortunately, very few results have been obtained for DAL. Decidability as well as a Hilbert-style axiomatization of DAL are still open. In [12], a variant of DAL, called $DALLA$ is proposed. The language of $DALLA$ is the same as that of DAL, but the semantics is based on $DALLA$-structures (cf. Sect. 2.2). A sound and complete deductive system consisting of the following two axioms in addition to $S5$ axioms for the operators is given.

1. $[a \uplus b]\alpha \leftrightarrow [a]\alpha \wedge [b]\alpha$.
2. $[a \cap b]\alpha \leftrightarrow [a]\alpha \vee [b]\alpha$.

3.3 Information Structures and the Logic by Balbiani

As pointed out by Orłowska in [30], when we say that two objects are indistinguishable in an information system, we actually mean that these are indistinguishable not absolutely, but with respect to certain properties/attributes. Thus in the study of indiscernibility relations, it seems important to bring the attribute set also into the picture. But the logics discussed so far lack this feature. In order to achieve this, Orłowska proposed the notion of an information structure (cf. Sect. 2.2), but referred the axiomatization of a logic with semantics based on information structures as an open problem. Later, Balbiani gave a complete axiomatization of the set of wffs valid in every information structure, using the technique of *copying* introduced by Vakarelov [43]. In fact, in [2], complete axiomatizations of logics with semantics based on various types of structures with relative accessibility relations is presented. One of these is a logic for information structures (cf. [1]). This, as required, is a multi-modal logic with a modal operator $[P]$ for each $P \subseteq \mathcal{A}$. Apart from the $S5-$axioms for each modal operator, the axiom $[P]\alpha \vee [Q]\alpha \rightarrow [P \cup Q]\alpha$ is considered. The canonical model obtained for this system only satisfies the condition $R_{B\cup C} \subseteq R_B \cap R_C$. Such a model is called *decreasing*. Using the method of copying, one obtains from a decreasing model, a model that satisfies condition $R_{B\cup C} = R_B \cap R_C$ and preserves satisfiability as well.

3.4 Logic $LMSAS^D$ for multiple source systems

We note that the logics discussed above including epistemic logic $S5_n$ are not strong enough to talk about the notions of strong/weak approximations considered in [23] (cf. Sect. 2.2). Thus a quantified propositional modal logic $LMSAS^D$, different from modal logic with propositional quantifiers as well as modal predicate logic, is proposed in [15] with semantics based on $MSAS^D$s. The language has a set PV of propositional variables, and a set T of terms built with countable sets of constants and variables and a binary function symbol $*$. Formulae are got through the scheme: $p \mid \alpha \mid \alpha \wedge \beta \mid A\alpha \mid [t]\alpha \mid \forall x\alpha$, where $p \in PV$, $t \in T$, and A is the global modal operator. Thus quantification ranges over modalities. The semantics is defined with the help of a function which maps a term t to a finite subset of the set N of sources, $*$ being translated as union of sets. The function determines which equivalence relation is to be used to evaluate a modality involving a term t. A sound and complete axiomatization is obtained for $LMSAS^D$ consisting of the standard $S5$ and first order logic axioms, in addition to the following:

$$[t]\alpha \rightarrow [t']\alpha \text{ when } B(t) \subseteq B(t'),$$

where $B(t)$ denotes the set of all variables and constants used in the term t.

3.5 Logics with Attribute Expressions

The logics discussed so far lack an important aspect related to the study of information systems. The language of these logics cannot refer to attributes or attribute values which are essential parts of an information systems. In this section, we survey logics with attribute expressions.

Decision Logic: Decision logic (DL) is the simplest logic with this feature. It is a propositional logic, the language of which contains a set $\mathcal{A}$ of attribute constants and for each $a \in \mathcal{A}$, a finite set Val_a of attribute value constants. Using these constants, atomic wffs are formed which are of the form (a, v), $a \in \mathcal{A},\ v \in Val_a$, and are called *descriptors*. The wffs of DL are formed in the standard way using the descriptors and Boolean connectives $\neg, \wedge$. Semantics of DL is directly based on the information systems. The satisfiability of the wffs in an information system $\mathcal{S} := (U, \mathcal{A}, \{Val_a\}_{a\in\mathcal{A}}, f)$ at an object $x \in U$ is defined in the natural way. For instance,

$$\mathcal{S}, x \models (a, v) \text{ if and only if } f(x, a) = v.$$

A sound and complete deductive system for DL can be given consisting of the following axioms in addition to the propositional logic axioms.

1. $(a, v) \wedge (a, u) \leftrightarrow \bot$, for any $a \in \mathcal{A},\ u, v \in Val_a$ and $v \neq u$.
2. $\bigvee_{v\in Val_a}(a, v)$, for every $a \in \mathcal{A}$.

The proof of completeness theorem is very simple. We need to show that every consistent set is satisfiable. Consider the system $(W,\ \mathcal{A},\ \{Val_a\}_{a\in\mathcal{A}}, f)$, where W is the set all maximal consistent sets. Moreover, $f(w, a) = v$ if and only if $(a, v) \in w$. Axioms 1 and 2 guarantee that f is a total function. Now, one can prove $\alpha \in w$ if and only if α is satisfiable in the above information system. This gives us the desired result.

Dynamic Information Logic: A temporal dimension is added to the study of information system by Orłowska. In [27], the notion of an information system is extended by adding the concept of time. A set T of 'time points' and a linear order $<$ on T are included to define *dynamic information system* $\mathcal{DS} := (U, \mathcal{A}, \{Val_a\}_{a\in\mathcal{A}}, T, <, f)$, where $f : U \times T \times \mathcal{A} \to \cup\{Val_a : a \in \mathcal{A}\}$ is such that $f(x, t, a) \in Val_a$, for any $x \in U, t \in T, a \in \mathcal{A}$. A logic DIL for dynamic information systems is also proposed in [27]. In the language of DIL, atomic statements are descriptors of decision logic, together with an object constant x – so these are triples (x, a, v), and are intended to express: "object x assumes value v for attribute a". There are modal operators to reflect the relations $<$ and $<^{-1}$. So the set of wffs is defined following the scheme: $(x, a, v) \mid \neg\alpha \mid \alpha\wedge\beta \mid [<]\alpha \mid [<^{-1}]\alpha$. The truth of all statements of the language is evaluated in a model based on a dynamic information system, with respect to moments of time, i.e. members of the set T.

A DIL-model is a tuple $\mathcal{M} := (\mathcal{S}, m)$ where $\mathcal{S}$ is a dynamic information system, and m a meaning function which assigns objects, attributes and values from $U,\ A,\ Val$ to the respective constants. The satisfiability of a wff α in a model $\mathcal{M}$ at a moment $t(\in T)$ of time is defined inductively as follows:

$\mathcal{M},\ t \models (x, a, v)$ if and only if $f(m(x), t, m(a)) = m(v)$,
$\mathcal{M},\ t \models [R]\alpha$ if and only if for all $t' \in T$, if $(t, t') \in R$, then $\mathcal{M},\ t \models \alpha$,

$R \in \{<, <^{-1}\}$. For the Boolean cases, we have the usual definitions. A sound and complete deductive system of DIL can be given consisting of the axioms of linear time temporal logic along with the axiom,

$$(x, a, v) \wedge (x, a, u) \wedge (x', a, v) \rightarrow (x', a, u),$$

which says that the values of attributes are uniquely assigned to objects.

Logic *NIL*: The logic NIL (non-deterministic information logic) proposed by Orłowska and Pawlak [32] is an extension of the description logic by enriching the language with modal operators $\Box, \Box_1, \Box_2$ corresponding to similarity relation, inclusion relation and converse of inclusion relation. Wffs are built, as usual, out of the atomic wffs (descriptors) and the connectives. A NIL-model $\mathcal{M} := (U, S, R, m)$ consists of a NIL-structure (U, S, R) (cf. Sect. 2.2), along with a meaning function m from the set of all descriptors to the set 2^U. Satisfiability relation is defined in the usual way.

A sound and complete deductive system for NIL was proposed in [32], consisting of the following three axioms in addition to the KTB axioms for $\Box$ and $S4$ axioms for $\Box_1$ and $\Box_2$.

1. $\alpha \rightarrow \Box_1 \neg \Box_2 \neg \alpha$.
2. $\alpha \rightarrow \Box_2 \neg \Box_1 \neg \alpha$.
3. $\Box \alpha \rightarrow \Box_2 \Box \Box_1 \alpha$.

Note that due to Proposition 1, we also obtain completeness with respect to class of all standard NIL-structures.

Rauszer's Logic for Multi-agent Systems: Rauszer [39] describes a logic, that takes into account a (finite) collection of *agents* and their *knowledge bases*. We denote the logic as $\mathcal{L}_{\mathcal{MA}}$. The language of $\mathcal{L}_{\mathcal{MA}}$ has 'agent constants' along with two special constants 0,1. Binary operations $+, \cdot$ are provided to build the set $\mathcal{T}$ of *terms* from these constants. Wffs of a kind are obtained from terms, and are of the form $s \Rightarrow t,\ s, t \in \mathcal{T}$, where $\Rightarrow$ is a binary relational symbol. $s \Rightarrow t$ is to reflect that "the classification ability of agent t is at least as good as that of agent s".

Furthermore, there are attributes as well as attribute-value constants. Descriptors formed by these constants constitute atomic propositions, and using connectives $\wedge, \neg$ and modal operators $I_t,\ t \in \mathcal{T}$ (representing 'partial knowledge' of each agent), other wffs are formed.

$\mathcal{L}_{\mathcal{MA}}$-models are not approximation spaces, but what could be called 'partition spaces' on information systems. Informally put, a model consists of an information system $\mathcal{S} := (U, A, \{Val_a\}_{a\in A}, f)$, and a family of *partitions* $\{E_t\}_{t\in\mathcal{T}}$ on the domain U – each corresponding to the knowledge base of an agent. The family is shown to have a lattice structure, and the ordering involved gives the interpretation of the relational symbol $\Rightarrow$. Wffs built out of descriptors are interpreted in the standard way, in the information system $\mathcal{S}$. The partial knowledge operator I_t for a term t reflects the lower approximation operator with respect to the partition E_t on U. An axiomatization of $\mathcal{L}_{\mathcal{MA}}$ is presented, to give soundness and completeness results.

Some other Logics with Attribute Expressions: A class of logics with attribute expressions is also defined in [29, 28]. Models are based on structures of the form $(U, A, \{ind(P)\}_{P\subseteq A})$, where the 'indiscernibility' relation $ind(P)$ for each subset P of the attribute set A, has to satisfy certain conditions. For the models of one of the logics, for example, the following conditions are stipulated for $ind(P)$:

$(U1)$ $ind(P)$ is an equivalence relation on U,
$(U2)$ $ind(P \cup Q) = ind(P) \cap ind(Q)$,
$(U3)$ if $P \subseteq Q$ then $ind(Q) \subseteq ind(P)$, and
$(U4)$ $ind(\emptyset) = U \times U$.

Other logics may be obtained by changing some of $(U1)-(U4)$. The language of the logics has a set of variables each representing a set of attributes, as well as constants to represent all one element sets of attributes. Further, the language can express the result of (set-theoretic) operations on sets of attributes. The logics are multimodal – there is a modal operator to reflect the indiscernibility relation for each set of attributes as above. A usual Kripke-style semantics is given, and a number of valid wffs presented. However, as remarked in [28], we do not know of a complete axiomatization for such logics.

In the literature, one can find many other generalizations and extensions of DL without any hints on axiomatization, apart from those discussed above (eg. [10, 44, 9]). In [10], the features of arrow logic is added to define arrow decision logic. On the other hand, in [44], a generalized DL, denoted as GDL is proposed for interval set valued information system which itself is a generalization of information system where corresponding to each attribute a, objects are assigned a closed interval $[f_*(a), f^*(a)] := \{L \subseteq Val_a : f_*(a) \subseteq L \leq f^*(a)\}$ instead of an element of Val_a. If $f(x, a) = [f_*(a), f^*(a)]$, then the interpretation is that the object x definitely has properties in $f_*(a)$ and possibly has properties in $f^*(a)$. The language of GDL is same as that of DL but two types of satisfiability relations are defined.

$$x \models_* (a, v) \text{ if and only if } v \in f_*(a);$$
$$x \models^* (a, v) \text{ if and only if } v \in f^*(a).$$

Above satisfiability relations are extended to the set of all wffs in the standard way.

In [9], the notion of descriptors itself is generalized. There, descriptors are of the form (a, r, l) where,

a is an element of a finite set A of attribute symbols;

l is an element of a set L_a of label symbols;

$r \subseteq Val_a \times L_a$ is an element of a finite set R_a.

The semantics is based on the structure of the form $\mathcal{S}^+ := (\mathcal{S}, \{L_a : a \in A\})$, where $\mathcal{S} := (U, A, \{Val_a\}_{a \in A}, f)$ is an information system. Thus we have $x \models (a, r, l)$ if and only if $(f(x, a), l) \in r$. It is extended to the set of all wffs in the standard way. Note that, we obtain DL if we take $R_a := \{=\}$ and $L_a := Val_a$, for each $a \in A$.

Logic *LIS* for Information Systems: From the preceding discussions, one sees that a logic for information systems with all of the following features is yet to be obtained.

1. The language includes attribute and attribute value constants.
2. The semantics is based on a structure having relative accessibility relations with the power set of the set $\mathcal{A}$ of attributes as the parameter set. These relations in the structure are represented syntactically as modalities.
3. The relationship of the indiscernibility relation with the attributes and attribute values is reflected syntactically in the relationship between the modalities and the pairs of attribute-value constants.
4. There is a sound and complete deductive system for the logic.

In [21], a logic LIS for information systems is proposed having all the above mentioned properties. The language of LIS contains a finite set $\mathcal{A}$ of attributes, and for each $a \in \mathcal{A}$, a finite set $\mathcal{V}_a$ of attribute values. The wffs of the logic is given by the scheme: $(a, v) \mid p \mid \neg\alpha \mid \alpha \wedge \beta \mid [B]\alpha$, where $a \in \mathcal{A}$, $v \in \mathcal{V}_a$, $B \subseteq \mathcal{A}$, and $p \in PV$ is a propositional variable. A LIS-model is a tuple $\mathfrak{M} := (\mathcal{S}, V)$, where $\mathcal{S} := (U, \mathcal{A}, \{\mathcal{V}_a\}_{a \in \mathcal{A}}, f)$ is an information system, and $V : PV \to 2^U$.

The satisfiability of a modal formula $[B]\alpha$ is defined as,

$$\mathfrak{M}, w \models [B]\alpha \text{ if and only if for all } w' \in U \text{ with } (w, w') \in Ind_{\mathcal{S}}(B),\ \mathfrak{M}, w' \models \alpha.$$

For the other cases, we have the usual definitions.

A sound and complete deductive system for LIS is given consisting of the following axioms:

Ax1. All axioms of classical propositional logic (PL).
Ax2. $[B](\alpha \to \beta) \to ([B]\alpha \to [B]\beta)$.
Ax3. $[\emptyset]\alpha \to \alpha$.
Ax4. $\alpha \to [\emptyset]\langle\emptyset\rangle\alpha$.
Ax5. $\langle\emptyset\rangle\langle\emptyset\rangle\alpha \to \langle\emptyset\rangle\alpha$.
Ax6. $[C]\alpha \to [B]\alpha$ for $C \subseteq B \subseteq \mathcal{AC}$.
Ax7. $(a, v) \to \neg(a, v')$, for $v \neq v'$.
Ax8. $\bigvee_{v \in \mathcal{VC}_a}(a, v)$
Ax9. $(a, v) \to [a](a, v)$.

Ax10. $((b, v) \wedge [B \cup b]\alpha) \rightarrow [B]((b, v) \rightarrow \alpha)$.

In [16], the logic LIS is extended to obtain a dynamic logic for information systems, which can talk about flow of information and its effect on the approximations of sets. The presence of the properties 1-4 listed above in LIS plays a crucial role there.

3.6 Temporal Logic for Rough Set Theory

The logic DIL, discussed in Sect. 3.5, is the first proposal of temporal logic for rough set theory. Note that DIL does not have modal operators for indiscernibility relations induced by information systems. So DIL can express the changes in attribute values of the objects with time, but the language of this logic is not strong enough to talk about (changes in) set approximations. A temporal logic, called *temporal rough logic (TRL)* is proposed in [5] which overcomes this limitation of DIL. TRL can be viewed as a combination of temporal logic with operators $\oplus$ (next), $\ominus$ (previous), $\mathcal{U}$ (until) and $\mathcal{S}$ (since) and the logic of rough sets handling upper and lower approximations. The wffs of TRL are generated by the scheme: $p \in PV \mid \neg\alpha \mid \alpha \wedge \beta \mid \Box\alpha \mid \ominus\alpha \mid \oplus\alpha \mid \alpha\mathcal{U}\beta \mid \alpha\mathcal{S}\beta$. The semantics of TRL is based on a structure called *dynamic approximation space*, which is a finite sequence $\mathfrak{F} := \mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_N$, where $\mathcal{F}_i := (U, R_i)$ is an approximation space, $i = 1, 2, \ldots, N$. The elements of $T_{\mathfrak{F}} := \{1, 2, \ldots, N\}$ are the time points. Moreover, the relation R_t, $t \in T_{\mathfrak{F}}$ represents the information about the domain U of the objects at the time point t. A model for TRL is defined to be a tuple $(\mathfrak{F}, V)$, where $\mathfrak{F}$ is a dynamic approximation space, and $V : PV \rightarrow 2^U$. The satisfiability of a wff α in a model $\mathfrak{M}$ at $t \in T_{\mathfrak{F}}$ and $w \in U$, denoted as $\mathfrak{M}, t, w \models \alpha$, is defined inductively:

- For each propositional variable p, $\mathfrak{M}, t, w \models p$, if and only if $w \in V(p)$;
- The standard definitions for the Boolean cases;
- $\mathfrak{M}, t, w \models \Box\alpha$, if and only if $\mathfrak{M}, t, w' \models \alpha$ for all w' such that $(w, w') \in R_t$;
- $\mathfrak{M}, t, w \models \oplus\alpha$, if and only if $t > 1$ and $\mathfrak{M}, t-1, w \models \alpha$;
- $\mathfrak{M}, t, w \models \ominus X$, if and only if $t > 1$ and $\mathfrak{M}, t-1, w \models X$;
- $\mathfrak{M}, t, w \models \alpha\mathcal{U}\beta$, if and only if there exists r with $t \leq r \leq N$ such that $\mathfrak{M}, r, w \models \beta$, and for all k such that $t \leq k < r$, $\mathfrak{M}, k, w \models \alpha$;
- $\mathfrak{M}, t, w \models X\mathcal{S}Y$, if and only if there exists r with $1 \leq r \leq t$ such that $\mathfrak{M}, r, w \models Y$, and for all k such that $r < k \leq t$, $\mathfrak{M}, k, w \models X$.

Axoimatization of the logic TRL is cited as an open problem in [17], but a tableau based proof procedure for TRL is proposed in [14].

4 Conclusions

This article is an attempt to present some of the works done on the generalizations and logical aspects of rough set theory. The multiple-agent scenario is an important issue in artificial intelligence, and one would like to extend rough theory in such a situation. Although, the articles [37–39, 34, 40] study the rough set theory in multiple-agent scenario, the issue of counterparts of standard rough set concepts such as approximations of sets, definability of sets, membership function, is not addressed. This issue is raised first time in [17], and this study is continued in [18, 20, 19, 14, 23, 15, 24]. It is to be noted here that these studies are based on the assumption that each agent is equally preferred. But one may be interested in a situation where an agent is preferred over another agent of the system in deciding membership of an object. For instance, we could make the assumption that an agent will always prefer herself (i.e. her knowledge base) over the other agents of the system. Thus with this assumption, if we find that $x \in \underline{X}_{R_1} \cap B_{R_2}(X)$ and $y \in \underline{X}_{R_2} \cap B_{R_1}(X)$, R_1, R_2 being the knowledge bases of sources 1 and 2 respectively and $B_{R_i}(X) := \overline{X}_{R_i} \setminus \overline{X}_{R_i}$, then source 1 will put more possibility on x to be an element of X than y. Observe that in the above conclusion, not only the knowledge base of the sources but also the preference of source 1 is playing a role. In [22], a rough set model is proposed where a preference order on the set of agents is also considered. Moreover, notions of lower/upper approximations are proposed which depend on the knowledge base of the agents as well as on the position of the agents in the hierarchy giving the preference of agents.

Search for a suitable logic for information systems remains an issue for rough set community working on the logical aspects of rough set theory. In Section 3.5, we have seen that the logic LIS proposed in [21] has some nice features, which are desirable for any logic of information systems handling approximations of sets. In fact, presence of these features has made it possible to extend LIS to obtain a dynamic logic for information systems. This dynamic logic can be used to reason about the information flow and its effect on the approximations of sets. The logic LIS and its dynamic extension is extended in [16] to obtain corresponding logics for incomplete information systems where some of the information regarding attributes of the objects may be missing. It remains to be seen if one could extend these works to capture non-deterministic information systems as well where an object-attribute pair is mapped to a set of attribute values.

References

1. P. Balbiani. Axiomatization of logics based on Kripke models with relative accessibility relations. In E. Orłowska, editor, *Incomplete Information: Rough Set Analysis*, pages 553–578. Physica Verlag, Heidelberg New York, 1998.
2. P. Balbiani and E. Orłowska. A hierarchy of modal logics with relative accessibility relations. *Journal of Applied Non-Classical Logics*, 9(2-3):303–328, 1999.
3. M. Banerjee and M. K. Chakraborty. Rough consequence and rough algebra. In W. Ziarko, editor, *Proc. RSKD '93, Canada*, pages 196–207. Springer-Verlag.
4. M. Banerjee and M. A. Khan. Propositional logics from rough set theory. *Transactions on Rough Sets*, VI, LNCS 4374:1–25, 2007.
5. M. Banerjee and M. A. Khan. Rough set theory: a temporal logic view. In M. K. Chakraborty, B. Löwe, M. N. Mitra, and S. Sarukkai, editors, *Studies in Logic Vol. 15, Proc. Logic, Navya-Nyaya and Applications: Homage to Bimal Krishna Matilal, 3-6 Jan, 2007, Kolkata*, pages 1–20. College Publications, London, 2008.

6. S. Demri and E. Orłowska. *Incomplete Information: Structure, Inference, Complexity*. Springer, 2002.
7. D. Dubois and H. Prade. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17:191–200, 1990.
8. R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge*. The MIT Press, 1995.
9. T. F. Fan, W. C. Hu, and C. J Liau. Decision logics for knowledge representation in data mining. In *Proc. COMPSAC '01*, pages 626–631, Washington, DC, USA, 2001. IEEE Computer Society.
10. T. F. Fan, D. R. Liu, and G. H. Tzeng. Arrow decision logic. In D. Ślęzak, G. Wang, M. Szczuka, I. Düntsch, and Y. Yao, editors, *Proc. RSFDGrC '05, Regina, Canada*, volume 3641 of *LNAI*, pages 651–659. Springer Verlag, 2005.
11. L. Farinas Del Cerro and E. Orłowska. *DAL* – a logic for data analysis. *Theoretical Computer Science*, 36:251–264, 1985.
12. G. Gargov. Two completeness theorems in the logic for data analysis. Technical Report 581, Institute of Computer Science, Polish Academy of Sciences, Warsaw, 1986.
13. S. Greco, B. Matarazzo, and R. Słowiński. Rough membership and bayesian confirmation measures for parameterized rough sets. In D. Ślęzak, G. Wang, M. Szczuka, and I. Düntsch, editors, *Proc. RSFDGrC '05, Regina, Canada*, volume 3641 of *LNAI*, pages 314–324. Springer Verlag, 2005.
14. M. A. Khan. *Multiple-source Approximation Systems, Evolving Information Systems and Corresponding Logics: A Study in Rough Set Theory*. PhD thesis, Indian Institute of Technology Kanpur, 2010.
15. M. A. Khan and M. Banerjee. A logic for multiple-source approximation systems with distributed knowledge base. *Journal of Philosophical Logic*. In Press. Available online at http://www.springerlink.com/content/lk47276g5022mn08/.
16. M. A. Khan and M. Banerjee. Logics for information systems and their dynamic extensions. *ACM Transactions on Computational Logic*, 12(4). In Press. Available online at http://tocl.acm.org/accepted.html.
17. M. A. Khan and M. Banerjee. Formal reasoning with rough sets in multiple-source approximation systems. *International Journal of Approximate Reasoning*, 49(2):466–477, 2008.
18. M. A. Khan and M. Banerjee. Multiple-source approximation systems: membership functions and indiscernibility. In G. Wang, T. Li, J. W. Grzymała-Busse, D. Miao, and Y. Yao, editors, *Proc. RSKT'08, Chengdu, China*, volume 5009 of *LNAI*, pages 80–87. Springer-Verlag, 2008.
19. M. A. Khan and M. Banerjee. An algebraic semantics for the logic of multiple-source approximation systems. In H. Sakai, M. K. Chakraborty, A. E. Hassanien, D. Ślęzak, and W. Zhu, editors, *Proc. RSFDGrC'09, New Delhi, India*, volume 5908 of *LNAI*, pages 69–76. Springer-Verlag, 2009.
20. M. A. Khan and M. Banerjee. A dynamic logic for multi-agent partial knowledge information systems. In *Proc. LMSC'09 at ESSLLI'09, Bordeaux, France*, 2009. Available at http://www.irit.fr/ Andreas.Herzig/Esslli09/.
21. M. A. Khan and M. Banerjee. A logic for complete information systems. In C. Sossai and G. Chemello, editors, *Proc. ECSQARU'09, Verona, Italy*, volume 5590 of *LNAI*, pages 829–840. Springer-Verlag, 2009.
22. M. A. Khan and M. Banerjee. A preference-based multiple-source rough set model. In M. Szczuka, M. Kryszkiewicz, S. Ramanna, R. Jensen, and Q. Hu, editors, *Proc. RSCTC'10, Warsaw, Poland*, volume 6086 of *LNAI*, pages 247–256. Springer-Verlag, 2010.

23. M. A. Khan and M. Banerjee. A study of multiple-source approximation systems. *Transactions on Rough Sets*, XII, LNCS 6190:46–75, 2010.
24. M. A. Khan and M. Ma. A modal logic for multi-agent tolerance approximation systems. In M. Banerjee and A. Seth, editors, *Proc. ICLA'11, New Delhi, India*, volume 6521 of *LNCS*, pages 124 – 136. Springer-Verlag, 2011.
25. J. Komorowski, Z. Pawlak, L. Polkowski, and A. Skowron. Rough sets: a tutorial. In S. K. Pal and A. Skowron, editors, *Rough Fuzzy Hybridization: A New Trend in Decision-Making*, pages 3–98. Springer Verlag, Singapore, 1999.
26. T. Y. Lin and Y. Y. Yao. Neighborhoods system: measure, probability and belief functions. In *Proceedings of the 4th International Workshop on Rough Sets and Fuzzy Sets and Machine Discovery*, pages 202–207, November 1996.
27. E. Orłowska. Dynamic information system. *Fundamenta Informaticae*, 5:101–118, 1982.
28. E. Orłowska. Logic of indiscernibility relations. In G. Goos and J. Hartmanis, editors, *Proceedings of the Symposium on Computation Theory, Zabrów, 1984*, volume 208 of *LNCS*, pages 177–186. Springer Verlag, 1985.
29. E. Orłowska. Logic of nondeterministic information. *Studia Logica*, 1:91–100, 1985.
30. E. Orłowska. Kripke semantics for knowledge representation logics. *Studia Logica*, XLIX:255–272, 1990.
31. E. Orłowska and Z. Pawlak. Expressive power of knowledge representation systems. *International Journal of Man-Machine Studies*, 20(5):485–500, 1984.
32. E. Orłowska and Z. Pawlak. Representation of nondeterministic information. *Theoretical Computer Science*, 29:27–39, 1984.
33. P. Pagliani. Pretopologies and dynamic spaces. *Fundamenta Informaticae*, 59(2-3):221–239, 2004.
34. P. Pagliani. Information quanta and approximation spaces I: non-classical approximation operators. In *GrC '05: Proceedings of the 2005 IEEE Conference on Granular Computing*, pages 605–610. IEEE Press, 2005.
35. Z. Pawlak. *Rough Sets. Theoretical Aspects of Reasoning about Data.* Kluwer Academic Publishers, Dordrecht, 1991.
36. J. A. Pomykała. Approximation, similarity and rough constructions. ILLC prepublication series for computation and complexity theory CT-93-07, University of Amsderdam, 1993.
37. H. Rasiowa. Mechanical proof systems for logic of reaching consensus by groups of intelligent agents. *International Journal of Approximate Reasoning*, 5(4):415–432, 1991.
38. C. M. Rauszer. Knowledge representation systems for groups of agents. In J. Woleński, editor, *Philosophical Logic in Poland*, pages 217–238. Kluwer, Dordrecht, 1994.
39. C. M. Rauszer. Rough logic for multiagent systems. In M. Masuch and L. Polos, editors, *Knowledge Representation and Reasoning under Uncertainty*, volume 808 of *LNAI*, pages 161–181. Springer-Verlag, 1994.
40. A. Skowron. Approximate reasoning in MAS: Rough set approach. In *WI '06: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*, pages 12–18, Washington, DC, USA, 2006. IEEE Computer Society.
41. A. Skowron and J. Stepaniuk. Tolerance approximation spaces. *Fundamenta Informaticae*, 27:245–253, 1996.
42. D. Vakarelov. Abstract characterization of someknowledge representation systems and the logic NIL ofnondeterministic information. In Ph. Jorrand and V. Sgurev, editors, *Artificial Intelligence II*, pages 255–260. North–Holland, 1987.

43. D. Vakarelov. Modal logics for knowledge representation systems. *Theoretical Computer Science*, 90:433–456, 1991.
44. Y. Y. Yao and Q. Liu. A generalized decision logic in interval-set-valued information tables. In N. Zhong, A. Skowron, and S. Ohsuga, editors, *Proc. RSFDGrC '99*, volume 1711 of *LNAI*, pages 285–293, London, UK, 1999. Springer-Verlag.
45. W. Ziarko. Variable precision rough set model. *Journal of Computer and System Sciences*, 46:39–59, 1993.

Md. Aquil Khan
Fraunhofer Institute for Secure Information Technology SIT,
Darmstadt, Germany.
mdaquilkhan@gmail.com

Axioms for locality as product

Kamal Lodaya and R. Ramanujam

Abstract. We consider systems of finite state agents that are sequential in themselves and interact with each other by synchronously performing common actions together. We study reasoning about such systems along two lines. We consider an enrichment of Kleene's regular expressions with a parallel composition operator and offer a sound axiomatization of equality on expressions. We also study a simple linear time product temporal logic which is the standard one at a local level, and boolean combinations of located formulas at the global level. We offer a complete axiomatization of the valid formulas of this logic.

1 Introduction

When we study the dynamics of systems with concurrently evolving components that interact among themselves, a natural question arises: at what level of system description do we explicitly model concurrent dynamics? This is an old conundrum that has kept philosophers busy. One solution is to record concurrency at the atomic level, so that the basic system events are already comprised of several simultaneously occurring ones, and we study the temporal evolution of such snapshots. In this view, if three events occur simultaneously at an instant, and two in the next, there is no particular 'actor' that provides any continuity from one instant to the next. This model is close to the way physical sciences study phenomena and adopted by Petri nets [Pet].

An alternative viewpoint is to see concurrent action at a much higher level of description: complex processes that evolve concurrently and independently, but interact occasionally. In this view, the component processes are sequential (and hence exhibit no concurrency), and the entire system is merely a fixed parallel composition of these processes. Games in extensive form, especially those of partial information, are examples of such systems: the continuity provided by each player over time, articulated by the notion of strategy, is crucial for game evolution [OR]. In the theory of computation, this notion of actor or agent can be identified with a *location*: the concurrent components notionally describe a distributed system. This was a view pioneered by Edsger Dikstra [Dijk] and developed into a rich theory of *Process algebra* by Hoare [Hoa], Milner [Mil1] and others [BPS]. This is the viewpoint we discuss in this article.

In particular, we consider systems of emphfinite state agents that are sequential in themselves and interact with each other by synchronously performing common actions together. These common actions can be thought of as telephone calls: the caller waits for the other to pick up, they exchange information, end the call and proceed further on their own asynchronous way. This is as opposed to communication by mail where the sender does not wait for the recipient but proceeds asynchronously.

The semantics of such systems can be envisioned as follows: each agent, being sequential non-deterministic, can be seen as a tree and the parallel composition of such trees as generating a set of trees obtained by nondeterministic interleaving of the agent trees, subject to synchronizations. Process algebraists have studied this operator extensively.

Considered as automata, the parallel composition operator above corresponds to *synchronized product*, and on languages, the corresponding operator is that of *synchronized shuffle*. It is easily seen that these operations preserve regularity, in the sense that the product machines are again finite state and that synchronized shuffle of regular languages is regular.

Over such systems, we focus on one particular aspect: that of axiomatizations. A celebrated theorem of Kleene offers a syntax of rational expressions that precisely define regular languages, and a complete axiom system of equality over these expressions was provided by Salomaa [Sal]. In this context it is natural to look for equations over rational expressions that involve an operator for parallel composition as well. We earlier presented a complete axiomatization using a reduction of paralellism to interleaving [Lod] (as is common in process algebra), here we present a sound axiom system which does not adopt this reduction.

Equational reasoning of this kind can be considered *global reasoning*, in the sense that it is carried out by someone who observes the behaviour of the entire system as it evolves. An alternative is *local reasoning*, where we reason about each agent separately (as far as possible) and combine the properties in some systematic way to infer properties of the composite system. Such compositional reasoning is naturally formalized as inference rules in logics. This is the other approach taken in this paper: we study specifications of agent properties in *propositional temporal logic of linear time* PTL [MP] and their global combinations. Once again we present a axiom system, and prove its completeness.

There is an important technical motivation for such local presentations. In general, if we have m agents, each of which is a k-state machine, the global state space has k^m states. Such a blow-up, exponential in the number of agents, is referred to as the *state explosion problem*. On the other hand, if we can reason about each component by itself, we have only km states to navigate. Referred to as partial order based methods, these approaches tend to utilize the idea that the entire set of interleavings may be large, but working with representative interleavings may suffice for many interesting properties. Several tools were developed in the 1990's based on such intuition [GW,Val,Pel]. In recent times, with the advent of multi-core architectures and relaxed memory models, such methods are acquiring renewed importance.

What is interesting about these axiomatizations? When we consider top level parallelism we speak of parallel composition of sequential nondeterministic behaviours, and the central difficulty is that of determinizing the components separately, since local choices made by agents influences global choices at system level. This difficulty manifests in both the equational axiomatization and the inference rules for the temporal logic studied here.

A natural but challenging question relates to how far such techniques can be generalized when the number of interacting agents is not a fixed finite number, but unbounded, and hence potentially infinite. [RS] offers some tentative suggestions for such reasoning.

2 Languages and automata

Fix a finite set Σ as an *alphabet*. We will call its elements *actions*. A finite sequence of actions $w : \{1, \dots, n\} \rightarrow \Sigma$, such as $aabab$, is called a *word* over the alphabet. A set of words is called a *language*.

The syntax of **rational expressions** over Σ is given by:

$e \ ::= \ 0 \mid 1 \mid a, \ a \in \Sigma \mid e_1 e_2 \mid e_1 + e_2 \mid e_1^*$

The set of all words over Σ is Σ^* and the empty set is 0. The null word is written 1, we also use the same notation for the language $\{1\}$. In general, given expressions e_1, e_2 for languages L_1, L_2, the concatenation of their words is denoted $e_1 e_2$. The iteration e^* stands for the language formed by repeatedly concatenating words from e to form another word. For instance, given the language $\{aa, ab, b\}$, also written $aa + ab + b$, the word $aabab$ is in $\{aa, ab, b\}^*$, but the word ba is not. The null word 1 is always in any e^* (by taking words from e zero times).

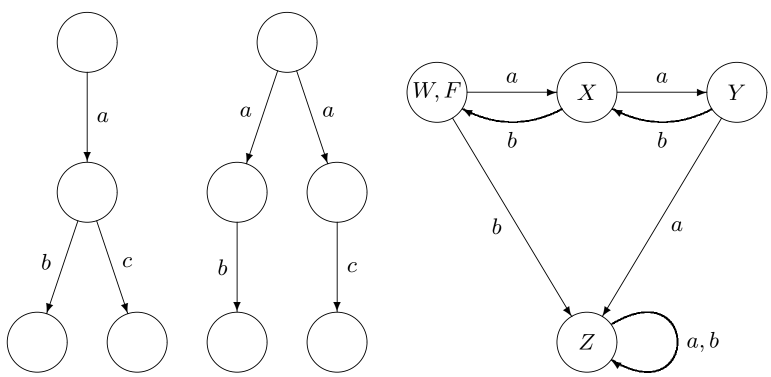


Fig. 1. The transition systems $a(b + c)$, $ab + ac$ and $Buff_2$

Definition 1. *A* ***labelled transition system*** *over the alphabet Σ is a directed graph $(Q, \rightarrow)$, with* states *Q and* transitions *$\rightarrow \subseteq Q \times \Sigma \times Q$. We will only be interested in transition systems which are* rooted *with a state $r \in Q$ from which all vertices are reachable. A* ***finite automaton*** *is a finite rooted transition system with a distinguished set of* final states*, let us say, marked by the variable F.*

We usually write $q \xrightarrow{a} q'$ to mean $(q, a, q') \in \rightarrow$, and interpret it to mean that the system can perform an action a at a state q and the resulting state is q'. A *run* of TS is a sequence $q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \ldots$ — a possible "execution" of the system. Labelled transition systems provide a natural model for the study of system behaviour. In Section 4 onwards we will only consider *maximal* infinite runs of a transition system.

A *run* of an automaton operating on a word begins in the root state. On each action, it takes the corresponding transition from the current state into a (possibly) new state. At the end of the word, if the automaton is in a final state, the run *accepts* the word. The *language accepted by the automaton* is all words for which there is a run from the root state to a final state. Two finite automata accepting the same language are said to be *equivalent*. The examples in Figure 1 show that they may be non-isomorphic as transition systems.

Suppose variables are used to denote states of a transition system. A state can be described in terms of the transitions going out to the other states. For example, the transition system $Buff_2$, which describes a language where the number of a's seen at any point in an accepting run can exceed the number of b's by at most 2, is given by the equations

$W = aX + bZ + F,\ X = aY + bW,\ Y = aZ + bX,\ Z = aZ + bZ$

For an expression e such as $aX + bZ$, its *initial actions* $Init(e) = \{a, b\}$, and X and Z are the a- and b-*derivative* of e respectively. The notation goes back to Brzozowski [Brz]. We say e has the *no empty word property* (NEWP) if the empty word is not in its language. This can be syntactically checked using derivatives.

2.1 Axiomatization

The Aanderaa-Salomaa axiomatization [And,Sal] for language equivalence of rational expressions is given below.

Axiomatization **RAX** for equivalence of rational expressions

(Assoc)	$(e + f) + g = e + (f + g)$
(Comm)	$e + f = f + e$
(Ident)	$e + 0 = e$
(Idemp)	$e + e = e$
(Assoc)	$(ef)g = e(fg)$
(Ident)	$e1 = 1e = e$
(Absorp)	$0 = 0e = e0$
(Distr)	$(e + f)g = eg + fg$
(Distr)	$e(f + g) = ef + eg$
(Guard)	$e^* = (1 + e)^*$
(Fixpt)	$e^* = 1 + ee^*$
(Fixpt)	$e^* = 1 + e^*e$
(GuardInd)	Let e have the NEWP. Then: $\dfrac{x = ex + f}{x = e^*f};\ \dfrac{x = xe + f}{x = fe^*}$

Theorem 1 (Salomaa). *The proof rules above are sound and complete for language equivalence of rational expressions.*

Proof. We only sketch the completeness. An inductive construction produces for any rational expression e a finite automaton accepting the language defined by e (for example, see [Koz]). Suppose e and f denote equivalent systems $TS(e)$ and $TS(f)$. By applying the axioms, each rational expression is reduced to a guarded sum of prefix form $\sum a_i e_i$ and $\sum b_i f_i$ respectively. Using left-distributivity, we need at most one derivative for each letter of the alphabet. Now equivalence between the roots guarantees equivalence between their successors, and equality among the derivatives of e and f guarantees derivability of $e = f$. So the task is reduced to proving completeness for nodes which are a distance 1 away from the roots. This can be repeated since all nodes of the transition systems are a finite distance away from the roots. $\square$

2.2 Solutions in rational expressions

Consider again the equations

$W = aX + bZ,\ X = aY + bW,\ Y = aZ + bX,\ Z = aZ + bZ.$

Using right-distributivity and introducing star, we get $Z = (a+b)^*$. Substituting for Z and then for Y, we get

$W = aX + b(a+b)^*,\ X = bW + a(a(a+b)^* + bX).$

Now we crucially need to apply left-distributivity. Following that up with another introduction of star, we have

$X = abX + bW + aa(a+b)^* = (ab)^*(bW + aa(a+b)^*).$

Applying the same medicine again,

$W = a(ab)^*bW + a(ab)^*aa(a+b)^* + b(a+b)^*$ and

$W = (a(ab)^*b)^*(a(ab)^*aa(a+b)^* + b(a+b)^*).$

This way of finding solutions is reminiscent of performing Gaussian elimination in linear arithmetic equations and was first used for regular languages by McNaughton and Yamada [MY].

Theorem 2 (Kleene). *The regular languages, those defined by rational expressions, are exactly those accepted by finite automata.*

Proof. We already referred to the forward direction in Theorem 1. Conversely, given a finite automaton TS, we apply the McNaughton-Yamada technique outlined above to end up with a solution which gives a rational expression for the root state. $\square$

3 Product words and product systems

Fix a nonempty finite set of *locations* $Loc = \{1, \ldots, n\}$. We now view the system alphabet $\Sigma \stackrel{\text{def}}{=} \Sigma_1 \cup \ldots \cup \Sigma_n$ as a *distributed* alphabet $\widetilde{\Sigma} = (\Sigma_1, \ldots, \Sigma_n)$, where each Σ_i is a finite nonempty set of *actions of agent* i. When an action a

is in $\Sigma_i \cap \Sigma_j, i \neq j$, we think of it as a *synchronization* action between i and j. (There can be k-way synchronizations also.) On the other hand, a *local* action is one in $\Sigma_i \setminus (\Sigma \setminus \Sigma_i)$, for some i.

We also make use of the associated implicit function $loc : \Sigma \to \wp(Loc) \setminus \{\emptyset\}$ which maps each action to the locations it is executed in, $loc(a) \stackrel{\text{def}}{=} \{a \mid a \in \Sigma_i\}$. For a set of actions A, $loc(A) \stackrel{\text{def}}{=} \bigcup\{loc(a) \mid a \in A\}$.

Given a distributed alphabet (Σ, loc), a *product word* is an element $(w_1, \ldots, w_n)$ of $(\Sigma^*)^{Loc}$ such that for some $w \in \Sigma^*$, every w_i is the restriction of w to actions in Σ_i. Thus $(w_1, \ldots, w_n)a$ is defined to be $(w'_1, \ldots, w'_n)$, where $w'_i = w_i a$ if $i \in loc(e)$ and $w'_i = w_i$ otherwise. If $loc(a) = \{i_1, \ldots, i_k\}$, we will call the expression $1||\ldots||a||\ldots||a||\ldots||1$, where a appears in positions $i_1, \ldots, i_k$ and 1 in the remaining positions, the distributed representation of a.

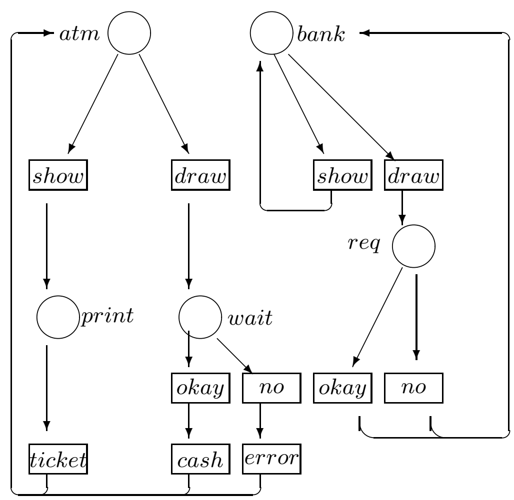


Fig. 2. A product system, the transition labels are boxed

We are interested in interactions between systems. A natural way to represent interactions among n agents is by having n transition systems, each working on its own alphabet of actions, except that the system undergoes "joint" transitions when common actions are encountered. Figure 2 models an ATM and a bank as a product system. A client at an ATM can ask for the balance to be shown and get a ticket printed. Alternately, the client asks to withdraw money and gets cash, or gets an error message, depending on whether the bank okays the transaction or not.

Definition 2. *Let $\widetilde{\Sigma} = (\Sigma_1, \ldots, \Sigma_n)$ be a distributed alphabet. A* ***parallel program*** *over $\widetilde{\Sigma}$ is a tuple $\mathcal{T} = (TS_1, \ldots, TS_n)$, where $TS_i = (Q_i, \to_i)$is a labelled* *transition system over Σ_i, for $i \in Loc$. When the individual transition systems are finite automata, we get a finite* ***product automaton*** *with final states* $\widetilde{F} \stackrel{\text{def}}{=} F_1 \times \cdots \times F_n$.

The global run is extended by an action a if and only if for every agent i participating in that action, a is enabled at the current local state of i. The obvious way to define global runs is to take products of transition systems.

Definition 3. *Let $\mathcal{T}$ be a parallel program over $\widetilde{\Sigma}$. The* ***product system*** *for $\mathcal{T}$ is the Σ-labelled transition system $TS = (Q, \Rightarrow)$, where*

- $\widetilde{Q} \stackrel{\text{def}}{=} Q_1 \times \cdots \times Q_n$, *and*
- $\Rightarrow \subseteq Q \times \Sigma \times Q$ *is the* ***global*** *transition function defined as follows:* $(q_1, \ldots, q_n) \stackrel{a}{\Rightarrow} (q_1', \ldots, q_n')$ *iff* $\forall i \in loc(a), q_i \xrightarrow{a}_i q_i'$, *and* $\forall j \notin loc(a), q_j = q_j'$.

We will use $q_1, q_2, \ldots$ to denote local states and $s_1, s_2, \ldots$ to denote global states. $s[i]$ will refer to the i^{th} component of the global state s. For the purposes of presentation we will assume a fixed distributed alphabet $\widetilde{\Sigma} = (\Sigma_1, \ldots, \Sigma_n)$ and the entire discussion will be relative to $\widetilde{\Sigma}$.

A product automaton operates on a product word $(w_1, \ldots, w_n)$ distributedly. Suppose it has inductively processed a prefix of this word reaching the global state $s = (s[1], \ldots, s[n])$, and $(u_1, \ldots, u_n)$ is the remaining suffix. Suppose that for all $i \in loc(a)$, all the u_i are of the form au_i' and for the remaining i, let $u_i' = u_i$. Then the corresponding transition is taken and the automaton moves to the new state tuple $s' = (s'[1], \ldots, s'[n])$, where for $j \notin loc(a)$, $s'[j] = s[j]$. The distributed representation of a is concatenated to the prefix and $(u_1', \ldots, u_n')$ is the remaining suffix. At the end of the word, if the automaton is in a final state in each component agent, the product word is accepted.

Suppose $\delta = s_0 \stackrel{a_1}{\Longrightarrow} s_1 \stackrel{a_2}{\Longrightarrow} \ldots$; by $\delta(k)$, we mean the global state s_k, whereas we use δ_k to denote the suffix of δ starting at s_k. We can meaningfully define a map $\lceil$, which, given a global run δ, and $i \in Loc$, retains only the i^{th} components of global states and erases all actions not in Σ_i. Clearly, $\delta \lceil i$ is a run of TS_i.

Further, suppose $\doteq$ is a binary relation on runs of $\mathcal{T}$, defined by: $\delta \doteq \delta'$ if and only if $\delta = \delta_1 s \stackrel{a}{\Rightarrow} s_1 \stackrel{b}{\Rightarrow} s_2 \delta_2$ and $\delta' = \delta_1 s \stackrel{b}{\Rightarrow} s_3 \stackrel{a}{\Rightarrow} s_2 \delta_2$, for some a, b such that $loc(a) \cap loc(b) = \emptyset$. Thus, δ and δ' are permutations of independent actions a and b. Let $\approx \stackrel{\text{def}}{=} (\doteq)^*$. The following assertion is easy to prove:

Proposition 1. $\delta \approx \delta'$ *iff for every* $i \in Loc, \delta \lceil i = \delta' \lceil i$.

Thus, we can think of δ as a representative of the equivalence class of δ under $\approx$ (denoted $[\delta]$), a non-sequential run of $\mathcal{T}$. This representation of product words is known in the literature as *Mazurkiewicz traces* [Maz]. Each trace over the distributed alphabet $\widetilde{\Sigma}$ can be thought of as the set of linearizations of the product word (the possible w's whose restrictions are w_i in the explanation above).

3.1 Parallel products of rational expressions

To model product systems over a fixed $\widetilde{\Sigma}$, we now introduce one outermost level for parallel product of rational expressions. This is much simpler than a language like Hoare's CSP [Hoa] which can have nested occurrences of the parallel operator.

$r^i \ ::= \ 0 \mid 1 \mid a, \ a \in \Sigma_i \mid r_1^i r_2^i \mid r_1^i + r_2^i \mid (r_1^i)^*$
$e \ ::= \ r^1 || \dots || r^n$

We generalize the Brzozowski derivative of a rational expression [Brz] to a distributed alphabet.

Definition 4. *$Der_a^{\widetilde{\Sigma}}(e_1 || \dots || e_n) \stackrel{\text{def}}{=} f_1 || \dots || f_n$, where for $1 \leq i \leq n$, if i in $loc(a)$ then $f_i = Der_a^{\Sigma_i}(e_i)$, otherwise $f_i = e_i$.*

This is an "expansion law" [Mil1], a global analysis seen as a product of actions on local components. Note that 1 appears in the derivative precisely when the local component can terminate, so termination of a product system is modelled by the expression $1 || \dots || 1$.

In reasoning about a parallel product we may have to identify the initial actions which will never make progress. Here is a sufficient condition.

Definition 5. *Let $\widetilde{\Sigma} = (\Sigma_1, \dots, \Sigma_n)$ and consider $e = e_1 + f_1 || \dots || e_n + f_n$. Suppose that for every i we have that $Init(e_i + f_i) \subseteq \Sigma_j$ for some $j \neq i$ and $Init(f_i)$ and $Init(f_j)$ are disjoint. We say that the actions in $Init(f_1), \dots, Init(f_n)$ (also the expressions $f_1, \dots, f_n$) are **useless in the sum** e.*

We now give a sound axiomatization for parallel products of rational expressions and an example of its use. The rules are parameterized by the arity of the parallel product, that is, the number of agents (n below) and depend on the distributed alphabet $\widetilde{\Sigma}$.

Axioms for equivalence of products **PAX=RAX+**

(Absorp) $0 = 0 || e = e || 0$

(Use) $e_1 + f_1 || \dots || e_n + f_n = e_1 || \dots || e_n$, if $f_1, \dots, f_n$ are useless in the sum.

(Deriv) $$\frac{Der_a^{\widetilde{\Sigma}}(e_1 || \dots || e_n) = Der_a^{\widetilde{\Sigma}}(f_1 || \dots || f_n), \text{ for all } a \in \Sigma}{e_1 || \dots || e_n = f_1 || \dots || f_n}$$

(ProdInd) Let $f_1, \dots, f_n$ have the NEWP. Then:
$$\frac{x_1 || \dots || x_n = (e_1 x_1 + f_1) || \dots || (e_n x_n + f_n)}{x_1 || \dots || x_n = e_1^* f_1 || \dots || e_n^* f_n}$$

The (Absorp) axiom models the fact that a deadlock in some part of the system means that the entire system is deadlocked. The (Use) axiom eliminates a useless chain of waiting. (By adding dummy sums like $\dots + a_i 0$ we may be able to eliminate useless chains of waiting through a subset of the n agents.) The (Deriv) rule was explained above. The (ProdInd) rule is a straightforward generalization of the (GuardInd) rule of **RAX** to the case of product systems.

3.2 Seeking solutions

To attempt a proof of completeness, we can apply the **RAX** axioms of Section 2.1 and reduce each product to guarded sum forms (from these the derivatives can be computed):

$$(\sum_{i_1} a_{i_1,1}e_{i_1,1})||\ldots||(\sum_{i_n} a_{i_n,n}e_{i_n,n}) \text{ and } (\sum_{j_1} b_{j_1,1}f_{j_1,1})||\ldots||(\sum_{j_n} b_{j_n,n}f_{j_n,n}).$$

We can further assume from the (Use) axiom that none of the initial actions is useless.

If one can proceed ahead using an action, we suppose that equality at the level of successors is derivable and use the derivative rule to conclude that $e = f$. Since overall there are finitely many possible derivatives for the expressions, either this strategy must succeed or we must come back to a situation seen earlier and the product induction rule can be used.

Here is a worked-out example. Using **RAX**, we get:
$X||Y = (a+ba)^*||(ab)^* = 1 + a(1+(a+ba)(a+ba)^*) + ba(a+ba)^*||1 + ab(ab)^*$.

Distributing and eliminating useless actions we have that:
$X||Y = 1 + aba(a+ba)^*||1 + ab(ab)^* = 1 + abaX||1 + abY = (aba)^*||(ab)^*$ by product induction.

To prove $X||Y = (aba)^*||(ab)^* = 1 + aba(aba)^*||1 + ab(ab)^*$ using derivatives, we will need to eventually show that:
$W||Z = aba(aba)^*||ab(ab)^* = ae||bf$ for some e, f. Eliminating useless actions, $W||Z = 0$. Hence $X||Y = 1||1$.

However, the weakness of the axiomatization is that it lacks a full analysis of all the cases which arise. In the next section we will see a temporal logic that uses another induction to solve this problem.

One can add a further axiom, Milner's expansion law, which reduces parallel product to interleaving (for example, $a||b = ab + ba$) and then directly uses the completeness of **RAX**. This route to establish completeness is explored in [Lod].

4 Temporal logic

Let $\mathcal{T} = (TS_1, \ldots, TS_n)$ be a parallel program, and let L_i denote the runs of TS_i, for $i \in Loc$, the local runs of agent i. The frames for our logic will be global runs, which represent arbitrary interleavings of actions of different agents. In accordance with verification literature on temporal logic (for example, [MP]) we will henceforth be concerned with infinite runs. Let $R_{\mathcal{T}}$ denote the set of all *maximal* runs of the product system for $\mathcal{T}$. We restrict our attention to only those programs which have at least one infinite run.

Definition 6. *A **frame** is a pair $F = (\mathcal{T}, \delta)$, where δ is an infinite run in $R_{\mathcal{T}}$.*

We now present the logical language which we will call $PrPTL$. Let $AP = \{p_0, p_1, \ldots\}$ be a countable set of atomic propositions with p ranging over AP. We use α, β, γ etc. (with or without subscripts) to denote local formulas. The syntax of **i-local formulas** is given by:

$\Phi_i \ ::= \ p \mid \neg\alpha \mid \alpha \vee \beta \mid \langle a \rangle_i \alpha \mid \alpha \mathbf{U}_i \beta$

where, $a \in \Sigma_i$. This is basically PTL, where the next state modality has been indexed by actions.

We let $\phi_1, \phi_2, \ldots$ range over global formulas, whose syntax is given by:

$\Phi \ ::= \ \alpha @ i, \alpha \in \Phi_i \mid \neg\phi \mid \phi_1 \vee \phi_2$

A *model* is a pair $M = (F, V)$, where $F = (\mathcal{T}, \delta)$ is a frame, and $V : Q \rightarrow \wp(AP)$ is the valuation function over Q, the set of all local states of the system. Thus, atomic propositions are evaluated at local states.

The formula ϕ being satisfied in a model M at a temporal instant is defined below. We first define the notion for i-local formulas over local runs in L_i. Let $M_i \stackrel{\text{def}}{=} ((\mathcal{T}, \delta\lceil i), V_i)$, where V_i is the restriction of V to Q_i. Let $\rho = \delta\lceil i \in L_i$.

- $M_i, 0 \models p$ iff $p \in V_i(\rho(0))$.
- $M_i, k \models \neg\alpha$ iff $M_i, k \not\models \alpha$.
- $M_i, k \models \alpha \vee \beta$ iff $M_i, k \models \alpha$ or $M_i, k \models \beta$.
- $M_i, k \models \langle a \rangle_i \alpha$ iff $\rho(k+1)$ exists, $\rho(k) \xrightarrow{a}_i \rho(k+1)$ and $M_i, k+1 \models \alpha$.
- $M_i, k \models \alpha \mathbf{U}_i \beta$ iff $\exists m \geq k$ such that $M_i, m \models \beta$, and for all $l : k \leq l < m, M_i, l \models \alpha$.

The derived connectives of propositional calculus such as $\wedge$, $\implies$ and $\equiv$ are defined in terms of $\neg$ and $\vee$ in the usual way. Let $True$ abbreviate the Φ_i-formula $p_0 \vee \neg p_0$ and let $False$ stand for $\neg True_i$.

The derived modalities $\Diamond$, $\Box$, $\bigcirc$ and $[a]_i$ are given by:

$\Diamond\alpha \stackrel{\text{def}}{=} True \mathbf{U}_i \alpha$; $\bigcirc_i \alpha \stackrel{\text{def}}{=} \bigvee_{a \in \Sigma_i} \langle a \rangle_i \alpha$.

$\Box\alpha \stackrel{\text{def}}{=} \neg\Diamond\neg\alpha$; $[a]_i \alpha \stackrel{\text{def}}{=} \neg\langle a \rangle_i \neg\alpha$; $\odot_i \alpha \stackrel{\text{def}}{=} \neg\bigcirc_i \neg\alpha$.

We now define the semantics of global formulas.

- $M \models \alpha @ i$ iff $M_i, 0 \models \alpha$.
- $M \models \neg\phi$ iff $M \not\models \phi$.
- $M \models \phi_1 \vee \phi_2$ iff $M \models \phi_1$ or $M \models \phi_2$.

We will use the notation $\widehat{a}$ to abbreviate the global formula $\bigwedge_{i \in loc(a)} (\langle a \rangle_i True) @ i$. We can use $\widehat{a}$ to denote enabling of action a. Note that for any model M, if $M \models \widehat{a}$, and $\delta(0) = s$, then there exists a global state s' such that $s \stackrel{a}{\Rightarrow} s'$.

The formula ϕ is **satisfiable** if $M \models \phi$ for some model $M = ((\mathcal{T}, \delta), V)$. ϕ is **valid** (denoted $\models \phi$) if ϕ is satisfied in every model M. On the other hand, for a formula $\alpha \in \Phi_i$, we say that α is *i-valid*, if for every model M, we have $M_i, 0 \models \alpha$.

The following proposition is trivial to prove, and is the basis for expecting procedures for verification of properties in PrPTL:

Proposition 2. *Let $M = ((\mathcal{T}, \delta), V)$ be a model and let $\delta \approx \delta'$. Let $M' = ((\mathcal{T}, \delta'), V)$. Then for every PrPTL formula ϕ, $M \models \phi$ iff $M' \models \phi$.*

5 The axiom system

We now present an axiomatization of the valid formulas. We have one axiom system for each agent in the system, and in addition a global axiom system to reason about synchronization. In some sense, this helps us isolate how much global reasoning is required.

AXIOMATIZATION **LAXi** FOR AGENT i

(A0i) All the substitutional instances of the tautologies of PC

(A1i) $[a]_i(\alpha \implies \beta) \implies ([a]_i\alpha \implies [a]_i\beta)$

(A2i) $\langle a\rangle_i True \implies [b]_i False, \quad a \neq b$

(A3i) $\langle a\rangle_i \alpha \implies [a]_i\alpha$

(A4i) $\alpha \mathbf{U}_i \beta \implies (\beta \vee (\alpha \wedge \bigodot_i \alpha \mathbf{U}_i \beta))$

(MPi) $\dfrac{\alpha, \quad \alpha \implies \beta}{\beta}$

(TGi) $\dfrac{\alpha}{[a]_i\alpha}$

The axioms are quite standard. (A2i) expresses the fact that in any run, the next move made by agent i is unique.

We use the notation $\vdash_i \alpha$ to mean that the formula $\alpha \in \Phi_i$ is a theorem of system **LAXi**. We will call α *i-consistent* when $\neg\alpha$ is not a theorem of **LAXi**.

Proposition 3. *Every theorem of **LAXi** is i-valid.*

Some remarks are in order, before we proceed to present the global axiom system. We haven't included any axioms (or rules) in the local axiom systems for eventuality. This is so because information about local reachability is quite useless in PrPTL. Even an apparently local specification as $\langle a\rangle_i \alpha$ is in reality a global eventuality specification when there are other agents $j \neq i, j \in loc(a)$.

This suggests that we need some reasoning about eventuality at the global level. In temporal logic, this is typically achieved by an induction axiom or rule. Unfortunately since we have only boolean formulas at the global level, we cannot expect an axiom. The standard form of temporal induction for reachability looks like this:

$$\frac{\text{Global Invariant} \implies \alpha \wedge \bigodot(\text{Global Invariant})}{\text{Global Invariant} \implies \Box\alpha}$$

Since no global next state modality is available, we can only hope for something like this in PrPTL:

$$\frac{\bigwedge_k \text{Local Invariant@}k \implies \alpha @ i \wedge \bigodot_j \text{ Local Invariant@}j}{\bigwedge_k \text{Local Invariant@}k \implies (\Box_i \alpha)@i}$$

We can in fact write sound rules in this form, but they are too weak to express global reachability. Note that the global invariant is to specify "being one of several reachable global states". Now consider two global states characterized by formulas α and β respectively. We can assume that they are of the form $\bigwedge_{i\in Loc} \alpha_i@i$ and $\bigwedge_{i\in Loc} \beta_i@i$. Now notice that the formula $\bigwedge_{i\in Loc} (\alpha_i \vee \beta_i)@i$ is only implied by $\alpha \vee \beta$, but does not imply it. Thus, combination of local invariants can in general specify global states which are not reachable, and we need to somehow specify the following:

$$\widehat{b} \wedge \bigwedge_{k} \text{Pre-move}@k \implies ([b]_j \text{ Post-move})@j, \text{ for } j \in loc(b),$$

$$\bigwedge_{k \notin loc(b)} \text{Pre-move}@k \wedge \bigwedge_{j \in loc(b)} \text{Post-move}@j \implies \text{Global Invariant}$$

and relate the global invariant to the local properties. Unfortunately, this turns out to necessitate an infinitary scheme.

Global axiomatization **GAX**

(A0) $(\neg\alpha)@i \equiv \neg\alpha@i$

(A1) $(\alpha \vee \beta)@i \equiv (\alpha@i \vee \beta@i)$

(A2) $\bigvee_{a \in \Sigma} \widehat{a}$

(MP) $\dfrac{\alpha, \quad \alpha \implies \beta}{\beta}$

(GG) $\dfrac{\vdash_i \alpha}{\alpha@i}$

(GM) $\dfrac{\bigwedge_{i \in loc(a)} \alpha_i@i \implies \bigvee_{j \notin loc(a)} \alpha_j@j}{\bigwedge_{i \in loc(a)} (\langle a \rangle_i \alpha_i)@i \implies \bigvee_{j \notin loc(a)} \alpha_j@j}$

Let $m > 0$ and $\alpha_1, \ldots, \alpha_m$ be formulas such that for all $l \in \{1, \ldots, m\}$, α_l is of the form $\bigwedge_{k \in Loc} \alpha_l(k)@k$. Let $\gamma \stackrel{\text{def}}{=} \bigvee_{l=1}^{m} \alpha_l$.

(Sy_m)
$$\dfrac{\begin{array}{l} \gamma \implies \neg\widehat{a} \\ \bigwedge_{l \in \{1,\ldots,m\}} (\alpha_l \implies (\bigwedge_{b \notin \Sigma_i} (\widehat{b} \implies \bigwedge_{j \in loc(b)} ([b]_j \beta(l,b,j))@j))) \\ \bigwedge_{l \in \{1,\ldots,m\}} \bigwedge_{b \notin \Sigma_i} ((\bigwedge_{k \notin loc(b)} \alpha_l(k)@k \wedge \bigwedge_{j \in loc(b)} \beta(l,b,j)@j) \implies \gamma) \end{array}}{\gamma \implies ([a]_i False)@i, \text{ for } i \in loc(a)}$$

(Un_m)
$$\dfrac{\begin{array}{l} \gamma \implies \neg\gamma_2@i \\ \bigwedge_{l \in \{1,\ldots,m\}} (\alpha_l \implies (\bigwedge_{b \in \Sigma} (\widehat{b} \implies \bigwedge_{j \in loc(b)} ([b]_j \beta(l,b,j))@j))) \\ \bigwedge_{l \in \{1,\ldots,m\}} \bigwedge_{b \in \Sigma} ((\bigwedge_{k \notin loc(b)} \alpha_l(k)@k \wedge \bigwedge_{j \in loc(b)} \beta(l,b,j)@j) \implies \gamma) \end{array}}{\gamma \implies \neg(\gamma_1 \mathbf{U}_i \gamma_2)@i}$$

The axiom (A2) ensures that some move is always enabled during the run. The rule (GG) allows us to globally infer theorems about agent i from those which have been proved in **LAXi**. (For instance, this rule, alongwith (A0) and (A1) allows us to infer "@-versions" of tautologies.) The rule (GM) specifies that when a global move labelled a is made, the local states of agents not involved in a remain unchanged. The (Sy) and (Un) rules describe eventual synchronization and the semantics of *until* formulas respectively.

$\vdash \phi$ is the notation used to denote theoremhood in **GAX**. ϕ is said to be *consistent* when $\neg\phi$ is not a theorem of **GAX**.

Proposition 4. *$\vdash \phi$ implies $\models \phi$.*

Proof. The axioms are obviously valid formulas. To show that the inference rules preserve validity, consider the rule (GM). Suppose that the premise is valid, but not the conculsion. Then there exists a model $M = ((\mathcal{T}, \delta), V)$ and $M \models \widehat{a}$, for every $i \in loc(a), M_i, 0 \models \langle a\rangle_i \alpha_i$, and for every $j \notin loc(a), M_j, 0 \models \neg\alpha_j$. We thus have $\delta(0) = s \overset{a}{\Rightarrow} s' = \delta(1)$, and for every $j \notin loc(a), s[j] = s'[j]$. It is easy to check that for every $i \in loc(a), M_i, 1 \models \alpha_i$, and for every $j \notin loc(a), M_j, 1 \models \neg\alpha_j$. Now consider the model $M' = ((\mathcal{T}, \delta'), V)$, where $\delta' = \delta_1$. Clearly, $M' \models \bigwedge_{i \in loc(a)} \alpha_i@i \wedge \bigwedge_{j \notin loc(a)} \neg\alpha_j@j$, contradicting validity of the premise.

Similarly, suppose that the premises of (Sy) are valid but that the negation of its conclusion is satisfiable. Then we have a model $M = ((\mathcal{T}, \delta), V)$ and $M \models \alpha_k@k$ for every $k \in Loc$. Further, $M_i, 0 \models \langle a\rangle_i True$, for some $i \in loc(a), a \in \Sigma$. then clearly there exist $b_1, \ldots, b_m$ such that $\{b_1, \ldots, b_m\} \cap \Sigma_i = \emptyset, \delta(0) \overset{b_1}{\Rightarrow} \delta(1) \ldots \overset{b_m}{\Rightarrow} \delta(m) \overset{a}{\Rightarrow} \delta(m+1)$. Now consider the models $M^l = ((\mathcal{T}, \delta_l), V)$, $l \in \{1, \ldots, m\}$. Obviously, since a is enabled at $\delta(m)$, we find that $M^m \models \widehat{a}$. Since the first premise says that $\gamma \implies \neg\widehat{a}$ is valid, it suffices to prove that $M^m \models \gamma$ to obtain a contradiction.

In fact, we argue that $M^l \models \gamma$ for every $l \in \{1, \ldots, m\}$. Firstly, since $M \models \gamma$ by assumption, for some $l, M \models \alpha_l$. Since $\delta(0) \overset{b_1}{\Rightarrow} \delta(1)$, we can show that $M \models \widehat{b_1}$, where $b_1 \notin \Sigma_i$. By validity of the second premise, for every $j \in loc(b), M \models ([b]_j \beta(l, b, j))@j$. therefore, we can show that for every $k \notin loc(b), M^1 \models \alpha_l(k)@k$ and for every $j \in loc(b), M^1 \models \beta(l, b, j)@j$. Thus, by validity of the third premise, we find that $M^1 \models \gamma$. We can repeat this argument to show that $M^l \models \gamma$, for $l > 1$ as well, and we are done.

The proof that the (Un) rule preserves validity is similar. □

Thus, we have soundness. We now proceed to show that **GAX** is indeed a complete axiomatization as well.

6 Completeness and decidability

In this section, we show that every consistent formula ϕ is satisfiable in a finite model whose size is bounded by $2^{c|\phi|}$. This at once shows that the logic is also decidable in nondeterministic exponential time.

We will find the following notation useful: when X is a finite set of formulas, by $\widehat{X}$ we mean the conjunction of all formulas in X (this is itself a formula).

Given Γ, a finite set, whose members are themselves finite sets of formulas, by $\widetilde{\Gamma}$, we mean the formula $\bigvee_{X \in \Gamma} \widehat{X}$.

We first define the notion of *subformula closure* of a formula.

Definition 7. *Let α be a formula in Φ_i.*

1. *$CL'_i(\alpha)$ is the least set of formulas containing α and satisfying the conditions*
   (a) *$\neg\beta \in CL'_i(\alpha)$ implies $\beta \in CL'_i(\alpha)$.*
   (b) *$\beta \vee \gamma \in CL'_i(\alpha)$ implies $\{\beta, \gamma\} \subseteq CL'_i(\alpha)$.*
   (c) *$\langle a \rangle_i \beta \in CL'_i(\alpha)$ implies $\{[a]_i\beta, \beta\} \subseteq CL'_i(\alpha)$.*
   (d) *$\langle a \rangle_i True \in CL'_i(\alpha)$, for every $a \in \Sigma_i$.*
   (e) *$\alpha \mathbf{U}_i \beta \in CL'_i(\alpha)$ implies $\{\alpha, \beta, \bigodot_i(\alpha \mathbf{U}_i \beta)\} \subseteq CL'_i(\alpha)$.*
2. *$CL_i(\alpha) \stackrel{\text{def}}{=} CL'_i(\alpha) \cup \{\neg\beta \mid \beta \in CL'_i(\alpha)\}$.*

For any α, $CL_i(\alpha)$ is finite and linear in the size of α. Similarly, we can define the subformula closure of global formulas.

Definition 8. *Let ϕ be a formula in Φ.*

1. *$CL'(\phi)$ is the least set of formulas containing ϕ and satisfying the conditions*
   (a) *$\alpha@i \in CL'(\phi)$ implies $\{\beta@i | \beta \in CL_i(\alpha)\} \subseteq CL'(\phi)$.*
   (b) *$\neg\phi_1 \in CL'(\phi)$ implies $\phi_1 \in CL'(\phi)$.*
   (c) *$\phi_1 \vee \phi_2 \in CL'(\phi)$ implies $\{\phi_1, \phi_2\} \subseteq CL'(\phi)$.*
2. *$CL(\phi) \stackrel{\text{def}}{=} CL'(\phi) \cup \{\neg\phi' \mid \phi' \in CL'(\phi)\}$.*

Once again, for any ϕ, $CL(\phi)$ is finite and linear in the size of ϕ. For convenience, we will abuse notation to also define $CL_i(\phi) \stackrel{\text{def}}{=} \{\beta | \beta@i \in CL(\phi)\}$, the set of i-subformulas of the global formula ϕ.

Fix a formula ϕ, and let $X \subseteq CL_i(\phi)$. We say that X is an *i-atom* of ϕ iff it satisfies the following conditions:

- for every $\beta \in CL_i(\phi), \neg\beta \in X$ iff $\beta \notin X$,
- for every $\beta \vee \gamma \in CL_i(\phi), \beta \vee \gamma \in X$ iff $\beta \in X$ or $\gamma \in X$,
- for every $a \in \Sigma_i$, if $\langle a \rangle_i True \in X$, then for every $b \in \Sigma_i, b \neq a, \langle b \rangle_i True \notin X$,
- for every $a \in \Sigma_i$, if $\langle a \rangle_i \alpha \in X$, then $[a]_i \alpha \in X$, and
- for every $\alpha \mathbf{U}_i \beta \in CL_i(\phi)$,
  $\alpha \mathbf{U}_i \beta \in X$ iff $(\beta \in X)$ or $(\alpha \in X$ and $\bigodot_i \alpha \mathbf{U}_i \beta \in X)$.

On the other hand, for a formula ϕ, we say that $A \subseteq CL(\phi)$ is a **atom** for ϕ iff for every $i, \{\alpha | \alpha@i \in A\}$ is an i-atom for ϕ. It can be easily checked that for every formula $\alpha@i \in CL(\phi)$, either that formula or its negation will be found in an atom (but not both). Further a formula of the form $\alpha@i \vee \beta@i$ (in $CL(\phi)$) is in an atom if and only if either of the disjuncts is in it. For an atom A, let $A[i]$ denote the associated i-atom, i.e. $\{\alpha | \alpha@i \in A\}$.

Let $AT_i(\phi)$ denote the set of all i-atoms for ϕ. Define $\rightarrow_i \subseteq AT_i \times \Sigma_i \times AT_i$ as follows: $X \xrightarrow{a}_i Y$ iff $([a]_i \alpha \in X$ implies $\alpha \in Y)$.

The global atom graph for a formula ϕ is defined as $G(\phi) \stackrel{\text{def}}{=} (AT(\phi), \Rightarrow')$, where $AT(\phi)$ is the set of all atoms for ϕ and $\Rightarrow'$ is defined by:

$$A \stackrel{a}{\Rightarrow}{}' B \text{ iff } \forall i \in loc(a), A[i] \xrightarrow{a}_i B[i] \text{ and } \forall j \notin loc(a), A[j] = B[j].$$

Suppose that $(W, \Rightarrow)$ is a subgraph of $(AT(\phi), \Rightarrow')$ in the sense that $W \subseteq AT$ and $\Rightarrow' \subseteq \Rightarrow$. Then we say that $(W, \Rightarrow)$ is *ϕ-good* if it satisfies the following conditions:

- there exists $A \in W$ such that $\phi \in A$, and
- for every $A \in W$,
  - A has a successor, i.e. B such that $A \stackrel{a}{\Longrightarrow} B$ for some a.
  - for every $a \in \Sigma$, if $\langle a \rangle_i True \in A$ for every $i \in loc(a)$, then there exists B such that $A \stackrel{a}{\Longrightarrow} B$.
  - for every $a \in \Sigma$, if $\langle a \rangle_i True \in A$ for some $i \in loc(a)$, then there exist $B_0, B_1, \ldots, B_k$, $k \geq 0$ and $b_1, \ldots, b_k$ such that $\{b_1, \ldots, b_k\} \cap \Sigma_i = \emptyset$, $A = B_0 \stackrel{b_1}{\Longrightarrow} \ldots \stackrel{b_k}{\Longrightarrow} B_k \stackrel{a}{\Longrightarrow} C$ for some C.
  - if $\alpha \mathbf{U}_i \beta \in A$ for some $i \in Loc$, then either $\beta @ i \in A$ or there exists B reachable from A such that $\beta @ i \in B$.

As it turns out, checking satisfiability of a formula ϕ amounts to checking the existence of such a ϕ-good subgraph in the syntactic graph $G(\phi)$. We will first show that every consistent formula does guarantee the existence of such a subgraph.

Lemma 1. *If ϕ_0 is a consistent formula, then $G(\phi_0)$ has a ϕ_0-good subgraph.*

Proof. Define W to be the set of maximal consistent subsets of $CL(\phi_0)$. (From now on, we fix ϕ_0 and use CL to mean $CL(\phi_0)$.) It is easy to check that each element of W is indeed an atom, using the local axiom systems, and (A0), (A1) and rule (GG). Thus $W \subseteq AT$. Simply define $\Rightarrow \stackrel{\text{def}}{=} \Rightarrow' \cap (W \times W)$. We claim that $(W, \Rightarrow)$ is ϕ_0-good.

Firstly, since ϕ_0 is consistent, there exists a maximal consistent set $A_0 \in W$ such that $\phi_0 \in A_0$. Now we have to prove that every element of W has an a-successor, and further that whenever $\{(\langle a \rangle_i True)@i | i \in loc(a)\} \subseteq A$, for $A \in W$, then A has an a-successor. Once we prove the second condition of these two, the first one follows, thanks to axiom (A2). Now, fix $A \in W$ and $a \in \Sigma$ such that $\{(\langle a \rangle_i True)@i | i \in loc(a)\} \subseteq A$. It can be easily checked that $\vdash \widehat{A} \implies \bigwedge_{i \in Loc} \widehat{A[i]}@i$. Further, let $X_i \stackrel{\text{def}}{=} \{\alpha | [a]_i \alpha \in A[i]\}$, for $i \in loc(a)$. Since $\langle a \rangle_i True \in A[i]$, we can show that $\vdash_i \widehat{A[i]} \implies \langle a \rangle_i \widehat{X_i}$. Thus, we find that $\vdash \widehat{A} \implies \bigwedge_{i \in loc(a)} (\langle a \rangle_i \widehat{X_i})@i \wedge \bigwedge_{j \notin loc(a)} \widehat{A[j]}$. But since $\widehat{A}$ is a consistent formula, so is the formula implied by it. then, by rule (GM), we find that $\bigwedge_{i \in loc(a)} \widehat{X_i}@i \wedge \bigwedge_{j \notin loc(a)} \widehat{A[j]}$ is consistent too. That is, the set $X \stackrel{\text{def}}{=} \bigcup_{i \in loc(a)} X_i \cup \bigcup_{j \notin loc(a)} A[j]$

is consistent. Hence, there exists a maximal consistent set $B \in W$ such that $X \subseteq B$. It can be easily checked that $A \overset{a}{\Rightarrow} B$ as required.

We now show that the eventual synchronization condition is also satisfied in the graph $(W, \Rightarrow)$. Fix $A \in W$ such that $(\langle a \rangle_i True)@i \in A$ for some $i \in loc(a)$. Let Γ be the least subset of W which satisfies the following conditions:

- $A \in \Gamma$, and
- whenever $B \in \Gamma$ and $B \overset{b}{\Rightarrow} C$ for any $b \notin \Sigma_i, C \in \Gamma$.

Note that every element in Γ is reachable from A via a path in $(W, \Rightarrow)$ which goes through actions outside Σ_i. Clearly, if there exists $B \in \Gamma$ such that B has an a-successor at all, then we are through.

Now suppose that there is no such B in Γ. We show that this assumption leads to a contradiction. It can be checked that $\vdash \widehat{B} \implies \neg\widehat{a}$ for every $B \in \Gamma$. Therefore, $\vdash \widetilde{\Gamma} \implies \neg\widehat{a}$.

Let $B \in \Gamma$. For every $b \notin \Sigma_i$, and for every $j \in loc(b)$, define

$$\Delta(B, b, j) \overset{\text{def}}{=} \begin{array}{ll} \{C[j] | B \overset{b}{\Rightarrow} C\}, & \textit{if} \vdash \widehat{B} \implies \widehat{b}, \textit{ and} \\ \{B[j]\}, & \textit{otherwise.} \end{array}$$

Claim (1). $\vdash \bigwedge_{B \in \Gamma} (\widehat{B} \implies (\bigwedge_{b \notin \Sigma_i} (\widehat{b} \implies \bigwedge_{j \in loc(b)} ([b]_j \widetilde{\Delta(B, b, j)})@j)))$.

Claim (2). $\vdash \bigwedge_{B \in \Gamma} \bigwedge_{b \notin \Sigma_i} (\bigwedge_{k \notin loc(b)} (\widehat{B[k]})@k \wedge \bigwedge_{j \in loc(b)} (\widetilde{\Delta(B, b, j)})@j) \implies \widetilde{\Gamma})$.

Suppose the claims are true. Then we have derived every premise in rule (Sy) in our axiom system. Hence, by its conclusion, $\widetilde{\Gamma} \implies ([a]_i False)@i$. But then $\widehat{A} \implies \widetilde{\Gamma}$ (since $A \in \Gamma$) and hence $\widehat{A} \implies ([a]_i False)@i$, clearly contradicting the fact that $(\langle a \rangle_i True)@i \in A$.

The only condition left to be proved (for $(W, \Rightarrow)$ to be ϕ_0-good) is the "until" condition. The proof of this proceeds in a fashion quite similar to the one for eventual synchronization, and hence is omitted here. ◻

Proof (of Claim (1)). Suppose that the formula in the Claim is not a theorem of the system. Then its negation is consistent, and we show that this leads to a contradiction. Skipping a few obvious steps, we can see that for some $B \in \Gamma, b \notin \Sigma_i$ and some $j \in loc(b)$, $\widehat{B} \wedge \widehat{b} \wedge (\langle b \rangle_j \neg \widetilde{\Delta(B, b, j)})@j$ is consistent. By the consistency of $\widehat{B} \wedge \widehat{b}$, we can expect B to have a b-successor. In addition, by a reasoning similar to what we employed earlier, we can find a b-successor, say C, such that $(\widehat{C[j]})@j \wedge (\neg \widetilde{\Delta(B, b, j)})@j$ is consistent. But, by construction of $\Delta(B, b, j)$, every such $C[j] \in \Delta(B, b, j)$. Then we get $\vdash (\widehat{C[j]})@j \implies (\widetilde{\Delta(B, b, j)})@j$, clearly a contradiction. ◻

Proof (of Claim (2)). This claim obviously follows from the fact that if $B \in \Gamma, b \notin \Sigma_i$ and $B \overset{b}{\Rightarrow} C$, then by construction, $C \in \Gamma$ and the fact that for $k \notin$

$loc(b), B[k] = C[k]$ by definition of $\Rightarrow$, whereas for $j \in loc(b), C[j] \in \Delta(B, b, j)$, again by construction. When B has no such b-successor, we get the required thesis by observing that $\widehat{B} \implies \Gamma$. □

Lemma 2. *If $G(\phi_0)$ has a ϕ_0-good subgraph, then ϕ_0 is satisfiable.*

Proof. Suppose that $(W, \Rightarrow)$ is a ϕ_0-good subgraph of $G(\phi_0)$. Let $A_0 \in W$ such that $\phi_0 \in A_0$. We claim that there is a maximal run δ of $(W, \Rightarrow)$ of the form $A_0 \overset{a_1}{\Rightarrow} A_1 \overset{a_2}{\Rightarrow} \dots$ which satisfies the following conditions: (let $k \geq 0$)

– If $(\langle a \rangle_i True)@i \in A_k$ then there exists $m \geq 0$ such that $A_{k+m} \overset{a}{\Rightarrow} A_{k+m+1}$ and for every l such that $0 \leq l < m$, $A_{k+l} \overset{b_l}{\Rightarrow} A_{k+l+1}$ implies that $b_l \notin \Sigma_i$.
– If $(\alpha \mathbf{U}_i \beta)@i \in A_k$, then there exists $m \geq 0$ such that $\beta@i \in A_{k+m}$.

The details of construction of δ are straightforward, though not trivial: we consider each of the n agents in a round-robin fashion, and keep fulfilling eventuality (until) requirements. Note that when an until-requirement is met for an agent, next-action requirements are also fulfilled upto the last action.

Now consider the parallel program $\mathcal{T} = (TS_1, \dots, TS_n)$ with $TS_i = (AT_i, \rightarrow_i)$, for $i \in Loc$. Let $TS \overset{\text{def}}{=} (Q, \Rightarrow')$ be the product system for $\mathcal{T}$. It can be checked that $\{(A[1], \dots, A[n]) | A \in W\} \subseteq Q$ and that $(A[1], \dots, A[n]) \overset{a}{\Rightarrow}{}' (B[1], \dots, B[n])$ iff $A \overset{a}{\Rightarrow} B$ in the given ϕ_0-good subgraph. Thus δ induces a maximal run δ' of the product system as well.

Now consider the frame $F = (\mathcal{T}, \delta')$, and the model $M = (F, V)$, where $V(X) \overset{\text{def}}{=} X \cap AP$, for $X \in AT_i$, for some i. Let $\rho = \delta' \lceil i$ and let $M_i \overset{\text{def}}{=} (\rho, V_i)$, where V_i is the restriction of V to AT_i.

Claim (3). For every $\alpha \in CL_i(\phi_0), k \geq 0, M_i, k \models \alpha$ iff $\alpha \in \rho(k)$.

Assuming the claim, we can go on to show that for all $\phi_1 \in CL(\phi_0), M \models \phi_1$ iff $\phi_1 \in A_0$. This is proved by an easy induction argument. But then, since $\phi_0 \in A_0$, it follows that $M \models \phi_0$, and we have demonstrated the satisfiability of ϕ_0. □

Proof (of Claim (3)). The proof proceeds by induction on the structure of α. The base case, when $\alpha \in AP$ is trivial and follows by the definition of V above.

The induction step proceeds by cases: when α is of the form $\neg\beta$ or of the form $\beta_1 \vee \beta_2$, the proof is by routine applications of the induction hypothesis. Now suppose that α of the form $\langle a \rangle_i \beta \in \rho(k)$. By construction of δ above (and hence of δ'), $\rho(k+1)$ exists, and $\rho(k) \overset{a}{\rightarrow}_i \rho(k+1)$. Further, since $\rho(k)$ is an i-atom, we find that $[a]_i \alpha \in \rho(k)$ as well, and by definition of $\Rightarrow$ above, $\alpha \in \rho(k+1)$. By induction hypothesis, $M_i, k+1 \models \alpha$ and hence $M_i, k \models \langle a \rangle_i \alpha$, as required.

On the other hand, when $M_i, k \models \langle a \rangle_i \alpha$, we are given that $\rho(k+1)$ exists, that $\rho(k) \overset{a}{\rightarrow}_i \rho(k+1)$ and that $M_i, k+1 \models \alpha$. By induction hypothesis, $\alpha \in \rho(k+1)$, and by definition of $\Rightarrow$, $\langle a \rangle_i \alpha \in \rho(k)$.

Now suppose α of the form $\beta \mathbf{U}_i \gamma \in \rho(k)$. If $\gamma \in \rho(k)$, we have (by induction hypothesis $M_i, k \models \gamma$ and hence $M_i, k \models \beta \mathbf{U}_i \gamma$. Otherwise, by construction of δ above, we find that there exists $m > k$ such that $\gamma \in \rho(m)$ and for all $l : k \leq l < m, \gamma \notin \rho(l)$. Now consider $\rho(k)$: since $\{\beta \mathbf{U}_i \gamma, \neg\gamma\} \subseteq \rho(k)$, being an i-atom, $\{\beta, \bigodot_i(\beta \mathbf{U}_i \gamma)\} \subseteq \rho(k)$. Hence $\beta \mathbf{U}_i \gamma \in \rho(k+1)$ as well. By the same reasoning $\beta \in \rho(k+1)$. Thus, we can show that for every $l : k \leq l < m, \beta \in \rho(l)$. The result follows by the induction hypothesis. The converse is proved similarly.

This completes the induction and the claim is proved. $\square$

The above two lemmas together lead us at once to the main results of the paper:

Theorem 3. *$\models \phi$ implies $\vdash \phi$; that is, **GAX** provides a complete axiomatization of the valid formulas of PrPTL.*

Theorem 4. *The satisfiability of a PrPTL formula ϕ can be decided by an algorithm taking $NTIME(2^{O(|\phi|)})$.*

We expect that the time complexity can be shown to be deterministic (singly) exponential time, using a more careful argument than the one presented above.

We can also consider the model checking problem for PrPTL: given a parallel program $\mathcal{T} = (TS_1, \ldots, TS_n)$, a PrPTL formula ϕ, and a valuation $V : Q \to 2^X$ (where X is the set of atomic propositions mentioned in ϕ), the problem is to determine whether every model based on $\mathcal{T}$ and V satisfies ϕ. By a standard argument, we can consider the product of the syntactic graph of ϕ above with the product system and check for connected components generating ϕ-good subgraphs. Here, we mention only the result:

Theorem 5. *Model checking a PrPTL formula ϕ against a parallel program with m global states is decidable in $NTIME(m.2^{O(|\phi|)})$.*

References

[And] Anderaa, S., On the algebra of regular expressions, in *Appl. Math. (course notes)*, Harvard, Jan 1965, 1–18.

[BPS] Bergstra, J., Ponse, A. and Smolka, S.A., eds., *Handbook of process algebra* Elsevier, 2001, 333–389.

[Brz] Brzozowski, J., Derivatives of regular expressions, *JACM* **11**,4, 1964, 481–494.

[Dijk] Dijkstra, E.W., *A discipline of programming* Prentice-Hall, 1976.

[GV] van Glabeek, R.J. and Vaandrager, F.W., Petri net models for algebraic theories of concurrency, *Proc. PARLE* **2**, Eindhoven (J.W. de Bakker, A.J. Nijman and P.C. Treleaven, eds.), *LNCS* **259**, 1987, 224–242.

[GW] Godefroid, P. and Wolper, P., A partial approach to model checking, *Inf. Comput.* **110**, 1994, 305–326.

[Hoa] Hoare, C.A.R., *Communicating sequential processes*, Prentice-Hall, 1985.

[Kle] Kleene, S.C., Representation of events in nerve nets and finite automata, in *Automata studies* (C.E. Shannon and J. McCarthy, eds.). Princeton, 1956, 3–41.

[Koz] Kozen, D., *Automata and computability*, Springer, 1997.

[LPRT] Lodaya, K., Parikh, R., Ramanujam, R. and Thiagarajan, P.S., A logical study of distributed transition systems, *Inf. Comput.* **119**,1, 1995, 91–115.

[Lod] Lodaya, K., Product automata and process algebra, *Proc 4th SEFM*, Pune (P.K. Pandya and D.v.Hung, eds.), IEEE, 2006, 128–136.

[MP] Manna, Z. and Pnueli, A., *The temporal logic of reactive and concurrent systems (Specification)*, Springer, 1991.

[Maz] Mazurkiewicz, A., Concurrent program schemes and their interpretations, Report DAIMI PB-78, Aarhus University, 1977.

[MY] McNaughton, R. and Yamada, H., Regular expressions and state graphs for automata, *IEEE Trans. Electr. Comp.* **9**, 1960, 39–47.

[Mil1] Milner, R., *A calculus of communicating systems*, *LNCS* **92**, 1980.

[Mil2] Milner, R., A complete inference system for a class of regular behaviours, *JCSS* **28**,3, 1984, 439–466.

[OR] Osborne, M.J. and Rubinstein, A., *A course in game theory* MIT Press, 1994.

[Pel] Peled, D., All from one and one from all: on model checking using representatives, *Proc CAV*, *LNCS* **697**, 1993, 409–423.

[Pet] Petri, C.-A., Fundamentals of a theory of asynchronous information flow, *Proc. IFIP*, Munich (C.M. Popplewell, ed.), North-Holland, 1962, 386–390.

[PW] Pinter, S., and Wolper, P., A temporal logic for reasoning about partially ordered computations, *Proc 3rd ACM PODC* , 1984, 28–37.

[Pnu] Pnueli, A., The temporal logic of programs, *Proc IEEE FOCS*, 1977, 46–57.

[RS] Ramanujam, R. and Sheerazuddin, S., Temporal specifications for services with unboundedly many active clients, to appear in *Journal of Applied Logics*, 2011.

[Sal] Salomaa, A., Two complete axiom systems for the algebra of regular events, *JACM* **13**,1, 1966, 158–169.

[Thi] Thiagarajan, P.S., A trace based extension of propositional linear time temporal logic, *Proc IEEE LICS*, 1994, 438–447.

[Val] Valmari, A., A stubborn attack on state explosion, *LNCS* **531**, 1990, 156–165.

Kamal Lodaya, R. Ramanujam
The Institute of Mathematical Sciences,
CIT Campus, Chennai 600113, India.
{kamal,jam}@imsc.res.in

জীবনের পথে দেখা হওয়া এক অনন্য ব্যক্তিত্ব

মাধবেন্দ্রনাথ মিত্র

কিছু কিছু মানুষ আছেন যাঁদের সান্নিধ্য মনের মধ্যে একটা ভাল লাগার অনুভূতি এনে দেয়। এই মানুষটির সংস্পর্শে যাঁরাই আসেন তাঁরা প্রায় সকলেই মনে করেন যে, এই মানুষটি অত্যন্ত কাছের মানুষ। এই রকম ভাল লাগার পিছনে কোন স্বার্থবোধ কাজ করে না। চৈত্র ব'লে এক ব্যক্তির মৈত্র নামে আর একজনকে ভাল লাগে, কারণ মৈত্রের সান্নিধ্যে এলেই তার কিছু লাভ হয় — হয়তো মৈত্র দেখা হলেই চৈত্রের প্রশংসা করে এবং তার আত্মমর্যাদাকে বাড়িয়ে দেয় অথবা চৈত্রের সঙ্গে দেখা হলে মৈত্র তাকে আর্থিক সাহায্য করে। এই সমস্ত ক্ষেত্রেও মৈত্রের সান্নিধ্য চৈত্রের ভাল লাগা সৃষ্টি করে। কিন্তু এই ভাললাগাটি কোন না কোন অর্থে স্বার্থজনিত। কিন্তু, এরকম কোন কারণে নয়, একদম স্বার্থগন্ধহীন অকারণে যে নির্ভেজাল ভাল লাগা, সেই ভাললাগার অনুভূতি দেয় যে মানুষটির সঙ্গ তাকে শ্রদ্ধা করতে ইচ্ছে করে, বিনা কারণে তার সান্নিধ্য ও সঙ্গ ভাল লাগে। এই ধরণের মানুষরা সাধারনতঃ আত্মপ্রশংসা করেন না, পরনিন্দা করেন না। এঁদের মধ্যে বিদ্বেষ ভাব থাকে না, এঁরা অহংকারী নন এবং কোন সময়ই ভান করেন না। এই রকম অকপটচিত্ত ঋজুস্বভাব মানুষকে আমরা সাধারনতঃ সজ্জন বলে থাকি। অধ্যাপক মিহির চক্রবর্তী আমার কাছে এরকম একজন সজ্জন ব্যক্তি। আমি অনুভব করি মিহিরের মধ্যে এক ধরনের স্বার্থগন্ধহীন ভালবাসা আছে, এবং তার ফলে অপরকে মর্যাদা দেবার একটা স্বাভাবিক প্রবণতা আছে। মিহিরের প্রতি আমার অকৃত্রিম শ্রদ্ধা ও ভাল লাগার গভীরে হয়ত এই সমস্ত কারণ বর্তমান।

আমার যতদূর মনে পড়ে মিহিরের সঙ্গে আমার প্রথম পরিচয় হয় ১৯৬৯ সালে। যে সময় মিহির বেলুড় কলেজে (রামকৃষ্ণমিশন বিদ্যামন্দির) গণিতের অধ্যাপক হিসেবে যোগদান করে। আমি তখন বিদ্যামন্দিরের দর্শনের অধ্যাপক। এই উজ্জ্বল এবং ঋজুপ্রকৃতির মানুষটির সঙ্গে কয়েকদিন মিশেই আমি তাঁর প্রতি আকৃষ্ট হয়ে পড়ি। কিছুদিনের মধ্যে বিদ্যামন্দিরের অন্যান্য শিক্ষকরা এবং ছাত্ররাও মিহিরকে সাগ্রহে গ্রহণ করে। আমরা সবাই জানি সত্তরের দশকে পশ্চিমবঙ্গে রাজনৈতিক বাতাবরণটি এমন ছিল যে, রাজনৈতিক মতান্তর প্রায়শঃই মনান্তরে পরিণত হয়ে যেত। ফলে একে অপরের প্রতি অসহিষ্ণু হয়ে পড়াটাই স্বাভাবিক বলে মনে হত। প্রায় সকলের মত মিহিরেরও একটি বিশেষ রাজনৈতিক মতামত ছিল। কিন্তু ও'কে কখনও অন্যের ব্যাপারে অসহিষ্ণু হতে দেখিনি। অন্যের সঙ্গে একমত না হয়েও তার সঙ্গে কি ভাবে আন্তরিক বন্ধুত্ব বজায় রাখা যায় — মিহিরের এই ইচ্ছা সকলেই অনুভব করত। অনেক বিষয়েই মিহিরের সঙ্গে আমার গভীর মতবিরোধ ছিল এবং আজও আছে, কিন্তু এই সমস্ত মতান্তর কখনই আমাদের পারস্পরিক আন্তরিক সর্ম্পককে কলুষিত করতে পারে নি। তাই আজও মিহির আমার একজন অকৃত্তিম সুহৃদ। আমি জানি, বিপদে, আপদে সম্পদে মিহির সর্বদাই আমার পাশে আছে এবং চিরকালই থাকবে। মিহিরের কারও সঙ্গে মতবিরোধ হলে তাকে কঠোর সমালোচনা করতে মোটেই পিছপাও হয় না, কিন্তু কখনই তাকে ছোট ক'রে মর্যাদাহানিকর ব্যবহার বা মন্তব্য করে না।

মিহির ক'লকাতা বিশ্ববিদ্যালয়ের বিশুদ্ধ গণিতের ছাত্র ছিল। গণিতের শিক্ষক হিসেবে সে বিদ্যামন্দিরে যোগদান করে। শিক্ষকতার প্রথম দিকে গণিত চর্চা করতে করতে ধীরে ধীরে মিহির লজিকের প্রতি আকৃষ্ট হয়ে পড়ে। বিশুদ্ধ গণিতের পাঠক্রমে লজিকের যেটুকু অংশ শেখানো হ'ত তা অতি সামান্য। পুরোপুরি লজিক চর্চা করতে হ'লে আরও অনেক ভালভাবে লজিক জানা প্রয়োজন। তাই লজিকের পড়াশুনোয় মিহির উৎসাহী হয়ে পড়ে। এই প্রসঙ্গে একটি মজার ঘটনা মনে পড়ে গেল। আমি তখন বিদ্যামন্দিরে সাম্মানিক দর্শনের ছাত্রদের ভারতীয় ও পাশ্চাত্য লজিক পড়াতাম। পাশ্চত্য লজিকে আবার দুটি প্রধান অংশ ছিল — একটিকে বলা যেতে পারে অবরোহ লজিক এবং অপরটিকে বলা হয় আরোহ লজিক। ইংরেজিতে যথাক্রমে এদের ডিডাকটিভ লজিক এবং ইনডাকটিভ লজিক বলা হয়ে থাকে। ডিডাকটিভ লজিক আবার দুভাগে পড়ানো হ'ত। তার মধ্যে একটি অংশ ছিল সাবেকি লজিক (ট্রাডিশনাল লজিক) এবং অপর অংশটি ছিল সাংকেতিক লজিক (সিমবলিক লজিক)। এই পাশ্চত্য লজিক পড়াবার সূত্রে মিহিরের আমার সঙ্গে বেশ কয়েকদিন সাবেকি ও সাংকেতিক লজিকের প্রাথমিক কিছু অংশ নিয়ে আলোচনা হয়। আমি খুব আগ্রহ ও উৎসাহ সহকারে মিহিরকে প্রাথমিক লজিকের পাঠ দিতে শুরু করি।

ক'য়েকদিন বাদেই আমি অনুভব করি মিহির আমার থেকে অনেক তাড়াতাড়ি লজিকের গভীরে ঢুকে যাচ্ছে। তখনই আমার এই সরল প্রচেষ্টাকে ধৃষ্টতা বলে মনে হতে শুরু করে। আজও মিহির অনেক সময়ই বিনয় সহকারে বলে থাকে যে ও'র লজিকের প্রথম পাঠ নাকি আমার কাছেই হয়েছে। কথাটা পুরো সত্য না হলেও আংশিক ভাবে সত্য। পাঠশালার গুরুমশাই ছাত্রদের অ, আ, ক, খ শেখান। তারপর সেই ছাত্ররাই পরবর্তী কালে কলেজ এবং বিশ্ববিদ্যালয়ের শিক্ষা সমাপ্ত ক'রে সেই গুরুমশাই-এর গুরু হবার যোগ্যতা অর্জন করেন, সেই অর্থে আমি মিহিরের লজিক শিক্ষার পাঠশালার গুরু। লজিকের বিষয় এখন মিহিরের যা পড়াশুনো ও বোধ, আমি চেষ্টা করলেও সেই গভীরে প্রবেশ করতে পারব না।

মিহির শুধু লজিকের পন্ডিত নয়, ও'র সাধারণ ভাবে সৃজনশীল মন লজিকের চর্চাকে অনেক বেশি উজ্জীবিত করে বলে আমার বিশ্বাস। এই প্রসঙ্গে একটি কথা বলতে চাই। আমার খুব আবছা আবছা মনে পড়ছে ১৯৮৫ সাল নাগাদ মিহিরের প্রচেষ্টায় কানাডার রিজাইনা বিশ্ববিদ্যালয়ের লজিকের অধ্যাপক আলফ্রেড টারস্কি'র প্রত্যক্ষ ছাত্র শ্রী হরগৌরী নারায়ণ গুপ্তের কয়েকটি বক্তৃতার আয়োজন হয়। সেই বক্তৃতা শুনতে লজিকের অনুরাগী বহু ছাত্রছাত্রী এবং বন্ধুবান্ধব ও সহকর্মী উপস্থিত হয়। এই বক্তৃতাগুলি এতই প্রাণবন্ত ছিল যে, সকলেই এর দ্বারা উদ্বুদ্ধ হয়। হরগৌরীবাবু চলে যাবার পরেও এই বক্তৃতার রেশ থেকে যায়। এই রেশকে কাজে লাগিয়ে মিহির একটি লজিক সার্কেল প্রতিষ্ঠা করার জন্য প্রস্তাব করে। মিহির ছাড়াও আরো কয়েকজন লজিকে অনুরাগী মানুষ এই প্রচেষ্টাকে সফলতা দান করে। এই কেন্দ্রটি আনুষ্ঠানিক ভাবে স্থাপিত হয় নি। তাই এর কোন আনুষ্ঠানিক রূপ নেই। অপরপক্ষে এটিকে একটি সম্পূর্ণ ইনফর্মাল প্রতিষ্ঠান বলা যেতে পারে। মজার বিষয় এই যে, এই লজিক চর্চার কেন্দ্রটি প্রায় পঁচিশ বছর ধরে পুরোদমে চলছে। আমি কচিৎ কদাচিৎ এই চর্চা কেন্দ্রের আলোচনা সভায় যোগ দিয়েছি এবং লক্ষ্য করেছি এই কেন্দ্রটির প্রাণপুরুষ হচ্ছে মিহির। আমার মনে হয়েছে, মিহিরের সান্নিধ্যের আকর্ষণই এই কেন্দ্রটিকে জীবন্ত রেখেছে। এইখানে যে সমস্ত আলোচনা হয়, তা আধুনিক লজিকের মূল স্রোতে থেকেও সৃজনশীলতাকে মূল্য দিয়ে থাকে। এই সৃজনশীলতার ফলে রক্ষণশীল দৃষ্টিভঙ্গিকে অতিক্রম করে লজিক যেন একটি উদারপন্থী দৃষ্টিভঙ্গি গ্রহণ করে। আমার বিশ্বাস মিহিরই এই উদারপন্থী ধারাকে এগিয়ে নিয়ে যেতে প্রশ্রয় ও উৎসাহ দেয়। এই প্রসঙ্গে বলা যেতে পারে মিহির বিশুদ্ধ লজিকের আলোচনায় সীমাবদ্ধ না থেকে লজিককে একটি নতুন মাত্রা দেবার কথা ভাবে। এই প্রচেষ্টাকে সফলতা প্রদান ক'রতে মিহির যেন বিশুদ্ধ লজিকের গন্ডী থেকে বেরিয়ে এসে লজিকের দার্শনিক দিকটার দিকে নজর দেবার কথা বলার চেষ্টা করে। প্রধানত মিহিরের উৎসাহের ফলে তাই বিশুদ্ধ লজিক চর্চা কেন্দ্রে আজকাল ভারতীয় দর্শনের লজিক চর্চাও ঢুকে পড়েছে, যাকে কোনক্রমেই বিশুদ্ধ লজিকের আওতায় ফেলা হয়ত সম্ভব নয়। এই দার্শনিক দৃষ্টিভঙ্গী এবং উদারপন্থী মনোভাব মিহিরের চিন্তাধারাকে একটি অনন্যতা প্রদান করেছে।

মিহিরের মূল বিষয় লজিক হলেও ও'র সৃষ্টিশীল মন লজিকের মধ্যে আবদ্ধ হয়ে যায় নি। মিহিরের সৃজনশীলতা প্রকাশ পেয়েছে তার সাহিত্যকর্মে। 'নান্দীমুখ' নামে একটি সাহিত্য পত্রিকার পেছনে মিহিরের অবদানকে কেউ অস্বীকার করতে পারবে না। তাছাড়াও বিভিন্ন সাহিত্য রচনা ও আলোচনার ক্ষেত্রে সাহিত্য জগতে মিহিরের অবদানকে আজকাল কেউই অস্বীকার করতে পারেন না। মিহিরের অধুনা লিখিত দুটি পুস্তিকা 'গণিতের ধারাপাত ও গল্পসল্প' এবং 'অমীমাংসার আলো-আঁধারি' পড়লে আমার কথাটির সত্যতা হয়ত উপলব্ধি করা যাবে। মিহিরের আরও একটি আগ্রহ ও প্রীতির বিষয় হছে চিত্র শিল্প। ও'র সৃজনশীল বোধের প্রকাশ বহু নাম করা চিত্র শিল্পীকেও মুগ্ধ করে। এছাড়াও মিহির মানবাধিকার আন্দোলনে আগ্রহী। শুধু মানুষ নয়, প্রকৃতি ও পরিবেশের অপব্যবহারও মিহিরকে উদ্বিগ্ন করে।

মিহিরকে যে কত লোক কত গভীর ভাবে ভালোবাসে সেটি আমি প্রথম বুঝেছিলাম তার একটি মরণাপন্ন ব্যাধির সময়। তখন আমি অবাক হয়ে দেখেছিলাম — কতলোক এই মানুষটির জন্য জীবনপণ করতে প্রস্তুত।

আর একটি কথা উল্লেখ না করলে মিহিরের জীবনের গল্পটি অসম্পূর্ণ থেকে যায়। সেটি মিহিরের স্ত্রী নমিতার কথা। নমিতা কবি। ও'র সঙ্গে মিশলে এবং ও'র কবিতা পড়লে বোঝা যায় ও'র ভেতরেও একটা ঔদার্য্যপূর্ণ্য ভালোবাসা ও'র ভেতরটাকে স্বচ্ছ করে রেখেছে। আমার মাঝে মাঝে মনে হয় মিহির কি মিহির হতে পারত যদি না নমিতা মিহিরের জীবনসঙ্গিনী হত। মিহিরের জীবনে যেমন শিল্প সাহিত্য কবিতা ও বিশুদ্ধ লজিক একাকার হয়ে গিয়েছে, ঠিক যেন তেমনই মিহির ও নমিতার জীবন একাকার হয়ে গেছে। ব্যক্তিস্বাতন্ত্রে বিশ্বাসী, নির্লোভ, উদার এবং সংবেদনশীল এই দম্পতি যেন একে অপরের পরিপূরক।

সবচেয়ে বড় কথা এদের দুজনের মধ্যেই কোথাও একটা ভালবাসা রয়েছে যা ব্যক্তিকেন্দ্রিকতায় আবদ্ধ নয়, যা অপর মানুষজনকেও কাছে টানে। মিহির ও নমিতা — যাকে বলে যোগ্যং যোগ্যেন যোজয়েৎ।

মিহিরের জীবনের সমস্ত দিকগুলি মনে রাখলে মানুষটাকে বুঝতে সুবিধা হবে। ঝকঝকে বুদ্ধি, দরদী মন, অনুভূতির গভীরতা, অন্যায়ের বিরুদ্ধে আপোষহীন মানসিকতা, ঋজুতা এবং সততা মানুষটিকে এমন একটা অনন্যতা দিয়েছে যে মানুষটিকে শ্রদ্ধা না ক'রে এবং ভাল না বেসে থাকা যায় না। এই রকম একটি সজ্জন মানুষ দীর্ঘজীবি হোন্ এই কামনা ক'রে আজ এখানেই শেষ করছি।

আমি মিহির সম্বন্ধে যে ক'টি কথা লিখলাম তা হয়ত তথ্যপ্রমাণ দিয়ে প্রতিষ্ঠা করা যায়, না গেলেও কিছু যায় আসে না। কারণ, এই ধারণা গুলি দীর্ঘ প্রায় চল্লিশ / বিয়াল্লিশ বছর ধরে আমার মনের মধ্যে গড়ে উঠেছে। কথাগুলি আমার অকৃত্রিম এক অনুভূতির স্বতঃস্ফূর্ত প্রকাশমাত্র।

Madhabendranath Mitra
Retired Faculty, Department of Philosophy,
Jadavpur University, Kolkata, India.
mnm_ju@yahoo.com

Can the feminists *qua* feminists be allies of M.C.?

Shefali Moitra

On many occasions Mihir Chakraborty (M.C.) and I have exchanged views on issues relating to abstraction. During these discussions, mine would be the stance of a feminist philosopher, and his would be the stance of a mathematician. M.C. has always been concerned with metamathematical questions. Many of these questions are relevant to the theoretical enterprise that preoccupies feminists as well. M.C. takes a somewhat non-conformist position in mathematics. For this reason he would be more acceptable to the radical feminist, than to the liberal feminist. At the outset let me clarify why I think M.C. is a non-conformist in mathematics. I shall then move to state in brief the difference between the liberal feminist and the radical feminist position in relation to theory construction. This should clearly bring out the relationship between M.C. on the one hand and the liberal and the radical feminist on the other.

According to the mainstream account of mathematics the subject is equated with rationality. The other name for mathematics is rationality and the other name for rationality is mathematics. M.C. characterizes the mainstream position by saying, 'for the mainstream "understanding means understanding through mathematics".' [1] Mathematics is often referred to as an exact science. One of the reasons why mathematics has been given a coveted position in the history of knowledge is that it is perceived as an instrument of success. (This continues to be the case in spite of the fact that predictability is not an epistemic virtue.) M.C. finds an echo of this perception in the words of William F. Taylor uttered during an interview in 1977; where Taylor says, 'quantitative science – that is science with mathematics – has proved successful in controlling nature. The majority of the society backs it for this reason. At the present moment, they want nature to be altered and controlled – to the extent, of course, that we can do it and the results are felicitous.' [2] The mainstream refers to mathematics as the queen of science. M.C. reminds us that this status is granted to mathematics as long as it delivers the desired results. Thus mathematics is not seen as an end-in-itself; all that it offers is an instrumental value. This is the majority view.

Things are however changing. Even Taylor in the above-mentioned interview goes on to say, 'the humanist point of view [regarding mathematics] is a minority point of view. But it is influential – one sees this among young people.' The assessment of mathematics as an exact science has been constantly changing over the past hundred years. M.C. quotes Bertrand Russell in his favour, 'mathematics is the subject in which we never know what we are talking about nor whether what we are saying is true.' [3] This is a recurring theme in M.C.'s writings. He keeps saying that the aim of mathematics is not to establish the truth. He does not believe that there is any such thing as the correct position in mathematics. He quotes Morris Kline who says, 'the major fact that emerges from the several conflicting approaches to mathematics is that there is no one body of mathematics but many.' [4] M.C. also denies the possibility of a unique logic. Till today many think that the Aristotelian laws of thought are sacrosanct. M.C. does not subscribe to this view. He quotes Rudolf Carnap in this context. Carnap says, 'in logic there are no morals. Everyone is at liberty to build up his own logic, his own form of language'.[5]

The acceptance of pluralism in logic and also in mathematics would be very close to the heart of the radical feminist. But the question is: who are the protagonists of this view? In other words does this view carry any authority? When it comes to the operative part we find that we are still guided by monolithic systems of logic and mathematics. In his article 'Jakhan Ar Kono Jijnasa Thakbe Na' M.C. wonders why this mainstream version of mathematics continues to prevail in the syllabus of the teaching institutions, especially at the primary level. M.C. envisages the urgency of looking into the social, cultural and commercial basis for this trend.[6]

M.C. repeatedly refers to the cultural and commercial influences on the selection of theories and on their dominance in the field of knowledge, including mathematics. In his 'Akashkusum-er Adhikar' he makes a passionate appeal. He says mathematics must free itself from the clutches of such dominance. This is the dominance of advertisements, the market, commerce and utility. Mathematics must return to its pristine realm of beauty. We should enter this realm with a kind of wonder 'mugdhata'.[7] We should not expect to receive a criterion for demarcating truth from falsity. That is not the task of mathematics. The relevant question to be asked is 'how did you like it?' (kemon lagche eke?). It is the 'beauty' of mathematics that is to be appreciated and enjoyed. Perhaps M.C. would be willing to use the word rasa in this context. Once we delve deep into the intricacies of mathematics we delve into its rasa as well. M.C. believes mathematics comes closer to poetry.[8]

The call to save mathematics from all forms of external influences and allow it to establish its autonomy is an attractive proposition especially for those of us who are mired in all sorts of power politics and are yearning to break away. It is like finding 'a room of my own'. The question is: has this ever been possible? Will this ever be possible? Or, is this a dream that the beleaguered mathematician likes to cherish? Did mathematics have a golden past when it had an intrinsic value? If so then what caused its enslavement, its cooption? M.C. gives a clue in one place. He says the eulogy showered on mathematics by the sciences especially physics has led mathematics to believe that it is the queen of sciences. It fails to realize that it is being praised for its 'use' value and not for its intrinsic value. M.C. says mathematics must come out of this orbit of sycophancy.[9] For him the question is not so much whether such autonomy was enjoyed by mathematics in the past or not. The question is whether men and women have the right to create a subject like mathematics in M.C.'s sense; is such a move legitimate? The question is one of adhikar. I would say of course one has the right to think of a situation where mathematics is interested in the creation of beauty alone, or a situation that accommodates mathematical pluralism. It is also plausible to approximate such a goal. But the possibility of the achievement of such a goal is problematic. Autonomy may be maximized but can absolute autonomy ever be achieved? The radical feminist would say the goal itself is a fantasy, some kind of absurdity or akashkusum.

The way M.C. argues in his article 'Akashkusumer Adhikar' it appears that the power-utility-market nexus is an avoidable evil.[10] Its influence on mathematics is to be seen as a fault-line which is to be got rid of. The liberal feminists would agree with him. They have a great faith in the autonomy of reason. They accept the fact that reason has been used for various vested interests, including power politics. For them the only way to escape from this unholy alliance between rationality on the one hand and power-market-utility on the other is to be more and more rational, more and more logical, more and more objective. The liberal feminists believe in the autonomy of reason, in the possibility of pure reason. For them a paradigm example of reason is to be found in Aristotelian logic. They believe in the existence of binaries such as reason/emotion, subjective/objective, logical/illogical.

Thus we find that though the liberal feminist would agree with M.C. on one major point (the point of escaping from a mundane context) they would not entirely endorse his position. Their point of agreement would be on the possibility of the autonomy of logic and mathematics. They would further state that human beings could achieve pure freedom only in the domain of pure rationality. Their allegiance to a rigid logical system would, however, distance them from M.C. They would not agree with him in ratifying Carnap's view stated above. M.C. agrees that there was a time when rationality had a pivotal role in emancipating man. Many say this about the Age of Enlightenment. Reason facilitated the rupture with the Dark Age of dogmatism. M.C. suggests the time has come now to free ourselves from the domination of reason. He does not intend to give up reason.[11] Not allowing reason to dominate implies the possibility of questioning reason itself.

By questioning rationalism alternatives to the mainstream dominant trend in reason may emerge. For instance, classical two-valued logic would demand a clear yes or no, true or false, right or wrong answer. M.C. does not subscribe to this requirement. He is a radical on this account. He says if he is asked which is true, the axiom of choice or its oppositional positions, he would prefer to remain silent on this issue not because he is undecided but because his answer will not be amenable to a clear-cut true or false categorization. He thinks the classical notion of truth was too arrogant – too self- confident, too self-promoting. Today we are willing to accept concepts like partial truth, mutually inconsistent truth, contingent truth, soft truth and so on.[12] This would sound like music to radical feminist ears. They are comfortable with a considerable amount of vagueness and ad hocism in theory construction.

The major difference between liberal feminism and radical feminism is that the liberals are pro-status quo. They are comfortable with the existing systems of science and justice. According to them what is at fault is the application of the system; therefore, the fault is extrinsic to the system. In places the fault lies in the non inclusion of women into the system. The liberal feminists primarily follow an agenda for inclusion into the mainstream. The radical feminists on the other hand feel that the existing systems are themselves faulty and intrinsically biased against women. What is needed is a systemic change.

It seems that the distinction between an extrinsic fault and an intrinsic fault of a system will be acceptable to M.C. He has spoken of the power-market-utility nexus that manipulates a system from outside. The intrinsic fault is caused by the arrogance of reason and its claim to exactitude. According to the radical feminist this fault is closely related to reason's denial of any association with feeling. For the radical feminist feeling is an inalienable part of woman's lived-experience. Therefore it should find its place in every human endeavour including theory construction. Since emotion carries an element of fuzziness the liberal feminist would like to keep reason and emotion in separate watertight compartments. I am not sure whether M.C. will allow emotion a constitutive role in mathematics. He talks about enjoyment in the encounter with mathematical systems, he also speaks of the beauty of the system, and these need not be constitutive components of mathematics. I think he may admit the role of intuition in mathematics and that would take him away from pure reason but it need not bring him close to emotion. It must be noted that M.C. grants an important role to imagination.

Moving away from reason does not necessarily imply some form of methodological anarchy. M.C. tells us so far whenever mathematics has changed it has replaced one consistent system by another consistent system. Metaphorically speaking one cage is introduced to replace another cage. For instance, Non-Euclidean geometry replaced Euclidean geometry. The choice between the two was always based on pragmatic considerations. The replacing of one cage by another is the procedure followed by mainstream mathematics, which M.C. does not subscribe to. He is not enamoured by a fully consistent mathematical system. He proposes a system that accommodates some amount of inconsistency or makes room for partial consistency. These proposed systems will have a constructed structure but the structure will not be flawless. M.C. extends the cage metaphor and says his cage may have some vulnerable bars, it may also have some broken bars and there will be gaps. There will be room for light as well as room for darkness in this cage. Clearly this alternative standpoint aims at avoiding hegemony by welcoming different voices and admitting that all is not crystal clear nor will it ever be.[13] Now we understand why M.C. stands against the mainstream position.

Having spoken in favour of extreme non-conformism in mathematics M.C. wonders why feminists insist on clubbing all abstractionists under one rubric. He wonders what could be the ground for holding all forms of abstraction responsible for gender bias. I had the privilege of receiving pre-publication comments from M.C. on my book *Feminist Thought* (2002). I admit that the feminist literature on the basis of which I made my observations in Feminist Thought were reacting to mainstream mathematics and logic and not to the non-conformist position adopted by M.C. and others. In the mainstream position it is believed that science has a universal mode of understanding based on mathematics. It is also believed that mathematics is based on logic, Aristotelian two-valued logic. It is further claimed that there is a unity between rationality, science and mathematics. There is no room for pluralism in this way of looking at abstraction. This point is explicitly elaborated by M.C.[14] Having followed the radical feminists I had made the statement that abstraction is a male-gendered virtue and, therefore, necessarily associated with authority and power.

M.C. had read the typescript and he commented that abstraction and formalism need not be associated with authority and power and, therefore, it could distance itself from all forms of gender politics. Abstraction could be a way to escape from mundane power politics. M.C. argued, authority is often very much part of our concrete lived-experience, it pressurizes us to the extent that we become numb. When the human condition gets so oppressive that we feel that there is no exit then the world of mathematics, logic and other forms of abstraction like music and art may offer some respite. Abstraction then offers an alternative to the coercion of our day-to-day concrete existence. I conceded to this point made by M.C. and added a note to the original text of *Feminist Thought*. In the note I wrote, 'The notions of the abstract and concrete are commonly contrasted in women's studies to establish the fact that abstraction is a male-gendered virtue, and therefore associated with authority and power. Such an association is a historical contingency. Power and authority could have been associated with an allegiance to the concrete. The dualism between the abstract and the concrete with the related privileging of one member of the pair must always be investigated for the detection of power games.' [15] This implies that all abstraction is not associated with oppression. Abstraction could at times be non-oppressive. Each case of abstraction has to be evaluated separately *in situ*.

Looking back at my book *Feminist Thought* almost after a decade I would like to formulate my response somewhat differently. The context of the remark 'abstraction is a male-gendered virtue and therefore necessarily associated with authority and power' needs clarification. In patriarchy the male-gender virtues are generally paired in the following manner so that the former member of the pair stands for male-gender virtues and the latter for female-gender virtues. The list commonly includes the following pairs of virtues: reason/emotion, abstract/concrete, assertive/submissive. Subsequently, the former characteristic of each pair is not only forwarded as a male-gender virtue, it is also touted as a human virtue. Since the former members of the pairs are foregrounded and the latter backgrounded a clear preference is indicated for male-gender virtues. The former virtues are associated with positive values and power and the latter members of the pair are associated with weakness and disvalue. I believe this gendered explanation of the relation between abstraction and power will be conceded by M.C. Here the relation between male-gender virtues and abstraction is a stipulated relation.

M.C.'s entry point and my entry point into the discussion of abstraction are very different. M.C. is looking at abstraction from the perspective of hope. The feminist and I are looking at abstraction from the standpoint of suspicion. If asked 'why is abstraction to be suspected?' The feminist will say that in the past abstraction has been projected as a guarantor of neutrality but on further analysis it has been shown that abstraction serves as a wonderful camouflage for partiality, more specifically for gender bias. The fact that various vested interests do perpetuate their hidden agenda under the garb of abstraction is a well-known fact. The liberal feminist will argue that these vested interests can be exposed given the time and effort, they can also be eliminated. In this way a biased abstract theory can be sanitized. This may be M.C.'s position as well. He admits the involvement of mathematics in power politics. In 'A Conversation with Helen Longino on Feminist Epistemology' published in Margins. M.C. remarks, 'set theory to me always appears to be a code of capitalist nations where only individuals and states remain, everything in between is wiped out. It seems to me very much political in that way.' [16] Why does M.C. say this? Is it not an admission of covert politics being encoded in theory? M.C.'s remarks on set theory seem to support the presence of politics that is intrinsic to mathematics. Earlier we had mentioned that M.C. accepts the political pressure on mathematics generated by the nexus of power, market and utility. This was an instance of external politics. Thus we find M.C. at times admits the presence of internal politics and at times accepts the presence of external politics in relation to mathematics. He also speaks of the possibility of extricating oneself from politics, both of the intrinsic type and the extrinsic type.

The feminists accept the possibility of the presence of gender politics at every level of cognition. Gender politics is unique because it causes a mind-set where gender discrimination is internalized. Subsequently the patriarchal mind-set is reflected in every individual action be it mental or physical. M.C. will perhaps agree with this possibility. The liberal feminist argues, with an effort one can transcend this mind-set and acquire a gender neutral/gender transcendent mind-set. They refer to this mind-set as the 'human' mind-set. The liberal feminist proposes to go beyond gender politics with the help of reason. In that case abstraction will facilitate the movement beyond mundane politics. For the liberal feminist reason is regimented and not loosely structured as M.C. would like it to be. The radical feminist does not accept the possibility of a future when gender will be transcended and a neutral context will be achieved. The radical feminist understands politics in terms of distribution of power. Power could be distributed in a top-down fashion as in patriarchy. This could be replaced by a feminist distribution of power; the structure could be one of power-sharing.

The liberal feminist agenda seems to be one of moving from patriarchy to some form of neutered humanism. In that case women's studies would only be a transitory phase to be transcended by humanism. For the radical feminist 'feminist thought' always has a political dimension. Politics is not necessarily bad. There can be good politics and there can be bad politics, but there cannot be any escape from it. Those who say they have got away from politics and are engaged in pure abstraction are also engaged with politics. The only difference is that the presence of politics is not openly admitted.

Does the presence of politics crash M.C.'s dream of unconditional free knowledge.[17] When politics is thrust upon a researcher then it is oppressive but when a researcher chooses a political paradigm it may not be oppressive. M.C.'s remark that mathematical activity is a part of human activity emboldens me to infer that he does not subscribe to the possibility of a view from nowhere.[18] Moreover, he clearly says that the mathematician has to create his or her own parameters, even though these parameters are temporary and relative.[19] M.C. agrees that the artist and the mathematician have to begin with some axiom. These axioms need not be derived from a factual basis, they come from convention or they come from imagination. Of these alternative possible grounds of an axiom imagination seems to be the most autonomous and the most unencumbered by context. To push the feminist critic a little further we may ask: can imagination be free from context, from political context? The radical feminist will say 'NO'. M.C. seems to have a similar proclivity. M.C. says in mathematics the days of Plato seem to have come to an end. Now we can state our preferences and decide which logic, which mathematics we would like to accept. Wittgenstein held that the growth of mathematics depends on our form of life. M.C. says he cannot give any final opinion on this issue, he acknowledges the importance of these debates.[20] The feminists take a stand in this ongoing debate. The liberal feminist think they can step out of the grid constituted by the 'form of life' and walk into a neutral zone, whereas the radical feminist believes that all our activities are circumscribed by some 'form of life'. I think M.C. would be more comfortable with radical feminists.

End Notes

1. M. Chakraborty: Akashkusumer Adhikar, in *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 13.

2. Ibid., p. 13-14.

3. M. Chakraborty: Sukumar Ray-er Logic in *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 73.

4. M. Chakraborty: Khela Bhangar Khela in *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 24.

5. M. Chakraborty: Sukumar Ray-er Logic in *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 75.

6. 'কেন যে পাঠশালাগুলিতে মূলত একটি ধারাই প্রচারিত এবং চর্চিত হয় তার সামাজিক, ব্যবসায়িক, সাংস্কৃতিক ভিত্তি খুঁজে বার করা দরকার বলে মনে হয়।' in M. Chakraborty: Jakhan Ar Kono Jijnasa Thakbe Na, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 8.

7. ‘বিজ্ঞাপণ, বাজার, ব্যবসা ও ব্যবহারের বাইরে তার নিজের, শুধুমাত্র সৌন্দর্য্যের কোণটিতে তাকে ফিরতেই হবে। ব্যবহারিকেরা আসুন সবিনয়ে, মুগ্ধতা নিয়ে, ভালবেসে।’ in M. Chakraborty: Akashkusumer Adhikar, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 14.

8. ‘ব্যবহারযোগ্যতা গণিতের কোন মূল ধর্ম নয়। ... বিজ্ঞাপণ-বাজার-ব্যবসা-ব্যবহারিকতার বাইরে শুধুমাত্র সৌন্দর্য্যের একান্ত কোণটিই তার নিজস্ব। এই অর্থে গণিত অনেক কাছাকাছি শিল্পের অথবা কবিতার। সুন্দরের যদি কোন সত্য থাকে সে সত্য গণিতেরও।’ in M. Chakraborty: Sundarer Satya, Ganiter Satya, *Amimangsar Alo-Andhari*, Nandimukh Samsad, Kolkata, 2010, p. 111-112.

9. ‘গণিতকে বেরিয়ে আসতেই হবে এই স্তাবকতার হাত থেকে, স্তাবকতার আড়ালে লুকোন লোভী দৃষ্টি থেকে।’ in M. Chakraborty: Akashkusumer Adhikar, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 14.

10. Ibid., p.14.

11. ‘জানি না কিছুদিনের মধ্যে বিজ্ঞান এবং র‍্যাশনালিজমও নতুন নতুন সমান্তরাল খাতে বইতে শুরু করবে কিনা। যে র‍্যাশনালিজম একদা মানুষকে সমষ্টিবদ্ধ সংস্কারের বাইরে আসতে সাহায্য করেছিল এখন সম্ভবত তার গন্ডি থেকেই বেরোতে চাইছে মানবমন।’ in M. Chakraborty: Kaler Jatrar Dhwani, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 64.

12. ‘আসলে সত্যাসত্যের এরকম একরোখা ধারণা হয়তো একটু দ্বিধায় ভুগছে আজকাল। প্রাচীন ‘সত্য’ একটু বেশি উদ্ধত ছিল মনে হয় - আত্মবিশ্বাসী এবং আত্মপ্রচারক ছিল। আংশিক সত্য, পরস্পরবিরোধী সত্য, পরিবর্তনীয় সত্য, কোমল সত্য ইত্যাদি কথাগুলি অবাঞ্ছিত নয় আজকাল আর।’ in M. Chakraborty: Ganitjagate Parbaparibartan, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 60.

13. ‘এতদিন গণিত যখন নিজেকে বদলেছে তখন একতি সুসংগত কাঠামোর পরিবর্তে আর একটি সুসংগত কাঠামো নিয়ে এসেছে, একটি খাঁচার বদলে আর একটি খাঁচা। ... খাঁচা আমাদেরও আছে। তা-ই আমাদের নির্মিতি, আমার বেঁচে থাকার প্রমাণ ... । তবে আমাদের খাঁচায় অনেক ফাঁক-ফোকর, ভাঙা শিক, পলকা বাঁধন, আলোর সাথে আঁধারও।’ in M. Chakraborty: Achin Pakhir Ghar, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 94.

14. ‘এই র‍্যাশনালিজম-এর সাথে অঙ্গাঙ্গী যুক্ত বিজ্ঞান, গণিত এবং সর্বোপরি লজিক। ... বিজ্ঞানের একটি সাধারণ চিন্তনপদ্ধতি আছে যার চূড়ান্ত চেহারাটি পাওয়া যায় গণিতে। আবার গণিতের ভিত্তিমূলে রয়েছে লজিক - দ্বি-মানযুক্ত অ্যারিস্টটলীয় লজিক (two valued logic) । ... এবং এই র‍্যাশনালিজম, বিজ্ঞান বা গণিত বলতে আমরা একটি মাত্র সমগ্র সত্তার কথাই বুঝেছি।’ in M. Chakraborty: Kaler Jatrar Dhwani, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 62.

15. Shefali Moitra: *Feminist Thought : Androcentrism, Objectivity and Communication*, Munshiram Manoharlal Publishers Pvt. Ltd., New Delhi, 2002, p. 144.

16. M. Chakraborty: A Conversation with Helen Longino on Feminist Epistemology, Margins, August 2000, p.11.

17. 'তাই যে যেমন খুশি এবং যখন খুশি ভাবতে আর বলতে পারার পরিবেশই গবেষণা, বা সৃষ্টিধর্মী যে কোন কাজের পূর্বশর্ত ... ('জ্ঞান যেথা মুক্ত')।' in M. Chakraborty: Khela Bhangar Khela, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 25.

18. 'গাণিতিক ক্রিয়াকলাপ তো মানবিক ক্রিয়াকলাপেরই অঙ্গ।' in M. Chakraborty: Khela Bhangar Khela, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 19.

19. 'অবশ্যই গণিত-কুশলী নিজেকে কোথাও বাঁধবেন, আপেক্ষিকভাবে, সাময়িকভাবে।' in M. Chakraborty: Achin Pakhir Ghar, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 93.

20. 'মনে হচ্ছে গণিতের ক্ষেত্রে প্লেটোর দিন শেষ হতে চলেছে। কোন নির্দিষ্ট একক গণিতজগতের অস্তিত্ব স্বীকার করা এবং সেই জগতের সত্যগুলি আবিষ্কারের জন্য সচেষ্ট হওয়া, এখন আর সম্ভব নয়। এখন বরং স্বাভাবিক মনে হয়, কোন গণিত আমার পছন্দ হয়, কোন লজিক আমি প্রয়োগ করব, কোন সেটতত্ত্ব আমি ভিত্তি হিসেবে নেব, এই বহুত্বধর্মী মনোভাব (approach) গ্রহণ। ব্যাক্তির পছন্দ-অপছন্দ নির্ভর হলেই বিষয়টি হয়ে পড়ছে সংস্কৃতি-নির্ভর। গণিত কি ব্যাক্তিমানুষের সংস্কৃতি-নির্ভর সামাজিক ক্রিয়া হতে যাচ্ছে আগামী দিন ? Wittgenstein বলেছেন, একটি সমগ্র জাতির রীতিনীতি, অভ্যাস এবং আচরণের উপর গণিতের বেড়ে ওঠা নির্ভর করে। ... আমি এই প্রসঙ্গগুলি সম্পর্কে কোন চূড়ান্ত মতামত দিতে পারব না এই মুহূর্তে।' in M. Chakraborty: Khela Bhangar Khela, *Ganiter Dharapat O Galpasalpa*, Nandimukh Samsad, Kolkata, 2001, p. 26.

Shefali Moitra
Retired Faculty, Department of Philosophy,
Jadavpur University, Kolkata, India.
smoitra22@rediffmail.com

Generalizing

Ranjan Mukhopadhyay

It is through philosophy that we seek a comprehension of the experiences we have. Indeed, it is a question of comprehension, because much of what we experience is baffling.

It is undeniable that we tend to achieve a coherence in what we experience along with its related reaction. This attempt at coherence brings in beliefs, sometimes conscious, sometimes unconscious. Consider for instance, although simplistically, the following series of experiences: feeling of pain in what seems to be a part of a body which is felt to be one's own, the visual experience of a squarish brownish edge protruding into that part of the body, the experience of an effort being made to dissociate the part of the body from the portruding edge, the experience of moving the part, the feeling of cessation of pain. This series is made coherent in holding that a part of my body strikes an object with an edge giving me pain and as a reaction I remove that part of my body from the vicinity of the object to give me relief from pain. The coherence achieved is seen to be both spontaneous and simultaneous with this series. The beliefs that are necessary for having this coherence may be said to be at least the following: that there is an I, a self, who has a body, that the self can feel pain, that the self can will to move its body, that such a will can get translated into a movement, that such a movement can cause cessation of pain.

The concepts of an I, or a self, of a body, of movement, of cause and effect, are inbuilt within the beliefs which seem to be unconscious in the sense that they are immediately given along with the experiences. That in this particular case moving a part of the body would result into, as an effect, the cessation of pain may be characterized as a belief that is conscious in the sense that it not immediately given with the experiences: this belief is formed and recalled from previous experiences.

The above example is not at all a happy one. Yet, it is relevant in that it helps one understand that certain concepts are coming into play in achieving a coherence within our experiences. What is an I, or a self? What is a body or an object? What is taken to be a movement? What is it for something to be a cause of something else? Or, conversely, what is it for something to be an effect of something else? These questions are the ones answers to which help us comprehend our experiences. Philosophy attempts to have answers to such questions.

It may be thought that, well, the concepts are given in any case as has been admitted above, and their importance has been stressed in their role of achieving coherence; fine, we understand that we cannot do without them, but what is this business about comprehension? What is the need for having a comprehension over and above achieving coherence? We feel this need when we ask: is the I, the self, an abiding one just because an assumption about its abidingness would be convenient for claiming the achievement of coherence? Or is it a series of I's, series of selves, as it is felt always to be in the present, and never felt to be in the past or in the future, although felt to have memories of the past moments? Which should be given more priority: conveniences of assumptions, or the felt experiences as foundations? Questions such as this baffle us.

Philosophy seeks comprehension in many modes. It can try to analyse a concept (in terms of other concepts already acceptable) in its quest for comprehension. It can try to relate a concept to others (which are, again, alredy available) for the same purpose. It can well try to abstract out general features of a concept for comprehension. There may be other ways of seeking comprehension as well, but we note, for the present, that the last one is a way in which a concept is attempted to be seen as one which is a special case of some other concept.

The example taken above may lead one to think that the concepts that come into play are all related to sensory experiences, and that philosophy is concerned with comprehension of only such concepts. On the contrary, examples are found in abundance where philosophy is concerned with concepts used in higher, or if we like other forms of experiences as well.

Take for instance the experience of coming to a conclusion for sure from a set of given premises. Concepts like premise, conclusion, inference come into play in achieving coherence regarding the experience. But the notion of a proposition, or a sentence, the notion of an argument, the notion of a psychological process of inferring, the notion of validity of arguments, etc. are introduced to have a comprehension of the concepts like premise, concluding for sure, etc.

Here too, it is found that the philosophical activity of comprehension of the notion of validity seems to have thrown up, in very rigorous and precise terms, a variety within itself. Validity in the context of the so-called truth-functional logic is simpler, in a sense, than that in the contexts of a so-called modal logic or the so-called first-order predicate logic. But the resulting, formidable looking and technical, notions seem to have an interconnection between them. It is found that the notion of validity in the context of truth-functional logic is a special case of the notion of validity in the context of a modal logic which again is a special case of the notion of validity in the context of a quantificational modal logic, which, as expected, includes the notion of validity in the context of the so-called first order predicate logic as a special case which again includes the case of the truth-functional logic as a special one.

This interconnection, in fact, is cast far wider in a topography of logics - and logics, than that instantiated above. The point to be noted is that comprehension in the mode of an abstraction or generalization is at play here. The interconnection in terms of generalization helps us in comprehending the layers existing in a notion brought in for giving an account of achieving coherence in our experiences.

In this mode of generalization the philosophical ingenuity lies in seeing the possibility of the existence of a layer (which has not been thought of earlier) at some point in a multy-layered notion. That uncertainty, arising from various directions, in the application of a predicate to a subject may call for the search of a suitable notion of validity (at a layer) that can be integrated in the multi-layered notion of validity was felt, among others, by Mihir Kumar Chakraborty. It seems that most of his works are directed toward such a search.

The concepts like those of a subject, a predicate, an application of a predicate to a subject and that of correctness of such an application are needed for the coherence of a kind of experience. At the level of comprehension the baffling questions that get asked are: is it that a predicate is such that either it applies to a subject or it does not? Or, is it that such applications are graded, that too seeming to be differently when the perspectives are changed? An affirmative answer, along with a rigorously and precisely attempted account of truth, to the first question shapes up a philosophical position. An affirmative answer to the second question in a similar fashion will let us have a different philosophical position. The debate about which of these positions have more justification than that of the other may go on indefinitely, or at least for quite some time in the future. But it will be an undisputed gain if it can be shown that a suitable notion of validity constructed on the basis of a notion of truth suitable for such graded applications of predicates to subjects can be integrated in the hierarchy of the multi-layered notion of validity. That such notions of validity for such contexts are available, or, have their places in the topography of logics gives a primary legitimacy to the notions themselves when the interconnections of these notions with the other more frequently and familiarly used notions of validity are established through the relation of one being the generalization of the other. This by itself is a valuable addition to the understanding of the notion of validity.

Chakraborty, I believe, has succeeded in establishing (in the above sense) the legitimacy of certain notions of validity for certain contexts which show some familiar and more frequently used notions of validity to be special cases of these. The technical works of his, unpacking these notions, have their own value. But from a philosophical point of view, his attempts at integrating the notions of validity that he excavates within the network of a layered notion of validity, done again through proofs, reveals an activity in the form of a mode of comprehension that has been caharacterized here as that of abstraction or generalization.

This aspect of his works may have been manifesting a personal trend of his. Metaphorically, Chakraborty, in terms of his views on uncertainty, is reluctant in commiting himself to hold on to a particular belief by taking recourse to the force of an analysis, he does not like to see himself placed permanently at some point in a network of beliefs either, but he is ready to have flights, in terms of generalizations, from one point to another all of which are really created by the flight itself. The flight and the course defining the flight indistinguishably merging in his case.

Ranjan Mukhopadhyay
Visva-Bharati,
Santiniketan, West Bengal, India.
ranjan.mukhopadhyay@visva-bharati.ac.in

Universality, locality, dialectic, dynamics and contradiction. Logico-political thoughts in honour of professor Mihir K. Chakraborty

Piero Pagliani

Abstract. By inspecting some questions in modern Logic, the paper tries to show how formal reasoning can be used as a guide to deal with socio-political problems, although it is not claimed that such problems can be actually formalised.

1 Formalism, social sciences and informal rigour

Since professor Mihir Chakraborty and I met for the first time, in Rome, some eighteen years ago, we discovered to share not only interests in logic and mathematics, but also in philosophy and, what is more important, in politics. As Marx used to say, Hindustan is an Italy at Asiatic dimensions. To me, this is true not only because of geographical similarity (both are peninsulas, with a long West to East river and a large plain in the north, the Ganga and the Po, surrounded by a bow of high mountains, the Himalaya and the Alps), but, also, from a political point of view. It is not by chance that an Italian patriot like Giuseppe Mazzini was greatly considered by Indian patriots. Moreover, after World War II, both nations experienced a long period of developmental state, a strong communist movement, and since the inception of liberalism in the Nineties, we both experimented a revival of nationalism, the Lega Nord in Italy and the BJP in India, together with a substantial bipartisan adoption of the neo-liberal doctrine.

As logicians with strong political interests, we discussed the possibility to set political analysis into a formal picture. However we have always been aware of the traps of what A. D. Sokal and J. Bricmon term "fashionable nonsense" or "intellectual imposture" (cf. [28]). They showed a plenty of amusing, irreverent and irritating examples of how formalism must not be used in human studies. Nevertheless, if one openly does not claim he is really formalising concepts from the social studies, then I maintain that comparing some social and cultural patterns with formal patterns, can help us to speak of culture and politics in a more controlled manner, so to avoid too much discretion in reasoning about those social phenomena.

In this paper I would like to give a taste of how some formal concepts can be used to understand the nature of capitalist social relationships. To do that, I shall mainly use notions from the very logico-mathematical field in which professor Mihir Chakraborty and I work together. It will be shown that some formal patterns and concepts help us to reason about capitalism and its contradictions in a more rigorous way, even if I do not claim that these patterns and concepts are able to model them in any formal way. Indeed our formal notions hardly can be maintained to be complete with respect to that universe of discourse. But, what is more important, they cannot even be considered consistent. Paradoxically, it is this very gap which makes the approach interesting, because, as we shall see at the end, it reflects the non coherence and non completeness of the capitalist system itself.

We can be aware of this fact only if we are able to give an interpretation of the impossibility to reduce social phenomena to formal models. To me, a first deep reason explained in the famous passage about Mathematics in Hegel's "Phenomenology of Spirit". This passage has been interpreted as an argument against the conceptual content of Mathematics. On the contrary, to me in that passage Hegel does not attack Mathematics in itself, but explains a real fact: formalism deals with the exposition of a mathematical result, not with its discovery, which, instead, is affected by a whole of intuition, visions, meditation, aesthetics and other non-formal mechanisms, as well as social interactions[1].

According to this interpretation, I maintain that understanding the real contradictions is a matter of Reason, while modeling them in a formal way is a matter of Intellect. Thus, any conceptual model, and particularly any formal model, cannot be but a partial approximation. Moreover, there is no way to extend it to a complete model, not because of formal limitations, but because of the gap between the means of the Intellect (which grasps just formal separations) and the means of Reason (which grasps real contradictions).

I start with some examples which show that also in a formalised field like Mathematical Logic, the idea of "one world, one logic", so that universe should have a ultimate coherent and complete logic (since there is a single ultimate universe), presents important problems. After that, examples are illustrated in order to provide the reader with some formal patterns of thought which can suggest schemes to interpret reality or, dually, to understand the distance between those schemes and real phenomena and the reason why there is that distance.

2 Combining specific logics of specific domains

Since the last decades of the XX century, it was found more and more urgent to combine the specific logics which arise from different theoretical objects and application fields, particularly Biology, Computer Science and Artificial Intelligence, mostly because of the new role played by the notion of "information". In a way, it was recognized that any specific object fulfills a specific logic (as claimed by Marx). Implicitly, the existence of an all-encompassing logic was excluded.

On the contrary, Capitalism claims that its own logic is an ultimate logic. But what is the logic of capitalism? Although the scientific paradigm of Marx's time was to proceed from an elementary seed (in Marx it was *the commodity*)

[1] For instance, the reading of previous authors or the need to explain something to someone else. The letter, for instance, was the motivation which pressed Dedekind to elaborate his construction of real numbers. Unlike Hegel, in my experience formalism is also a means to check the validity of the results of the above activities

which potentially embeds all the determinations of the reality to be described, and although Marx had his own powerful (and still valid, to me) logical and epistemological implementation of this method, sythesised by the concept of a "concrete abstraction" in the famous *1857 Introduction*, one may wonder whether Marx, using that method, exhibited the logic of actual capitalistic societies, or the logic of a *model* obtained by generalizing at an abstract level a particular form of capitalism, more precisely the one existing under the British global rule of the XIX century.

Probably the correct answer is "Both things" and it is the responsibility of Marx's follower to renew and apply the dialectic between theoretical model and real world interpretation.

Actually, in my opinion capitalistic contradictions break the overall logic of capitalism into different and often conflicting specific logics (for instance, social, political, economical and financial), while the capitalistic attempts to bypass those contradictions call for a manner to combine together the specific logics into which the capitalistic society is broken into fragments by its conflicting dynamics. On the other side of the barricade, it is the anti-capitalists' duty to understand how contradictions are managed by an alleged "total" logic, how these contradictions, on the contrary, break mutual relationships, how dominant capital tries to reassemble the fragments and, finally, how to intervene in, and exploit, this chaotic dynamics to abolish *the present state of things*.

We shall discuss the dialectic between the overall logic of capitalism and the contradictions induced by this logic in the last section. By now, let us note that combining different logics is an important research topic in applied formal logic, because there are situations which can be grasped by means of one logic, say a constructive one, while others require modal logic, and so on.

A very interesting approach is given by Dov Gabbay's Labelled Deductive Systems, LDS. In these systems, labels are used to control deduction. Any specific logic has its own specific labeling system (for instance in modal logics labels are possible worlds together with their accessibility relation, while in substructural logics labeling accounts for the occurrence and/or position of premises). These controls, are essentially provided by the (explicit or implicit) semantics of the different logics. Thus one can say that logics interact by means of the specific domains they refer to. Therefore, labels provide a means which is external to the deduction process. By means of this external control over the deduction steps, LDS provide a clear and appealing logical middleware which makes different systems interact (see [5]).

A different approach is given by Linear Logic, which provides a sort of "monadic system" encompassing both classical and constructive logics (see [6]). Semantics is definitely expelled by this approach, which is syntactic in nature, according to the assertion that the meaning of a formula is the set of its proofs. As a matter of fact, Classical Logic is the ultimate horizon of the construction. Such a unifying approach is more radical than the former, and appears as a system which, in a sense, is also a meta-system.

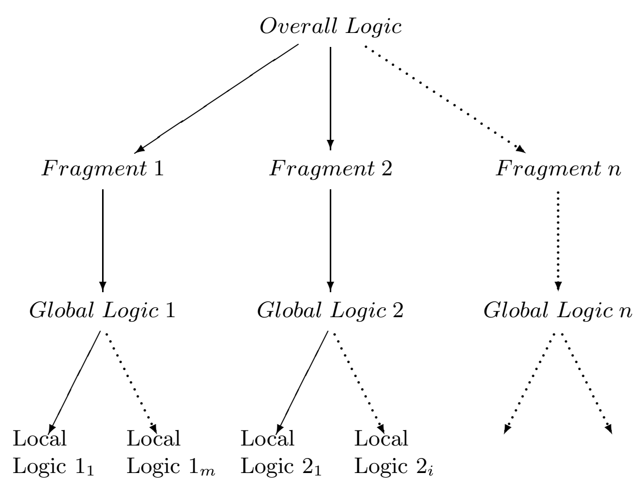
To synthesise, the former is a *pluralistic* approach whereas the latter is *monistic*. At this point, it is worth mentioning that also capitalist logic should be a system-and-metasystem, or code-and-metacode, whole because capitalism needs coherence in order to manage its terrific dynamics and conflicts. A coherence which should be guaranteed by the system itself, acting as a *self-justifying metasystem*. But because of those very dynamics and conflicts, such a coherence is continuously destroyed. Thus, reality deals with a whole which is continuously broken into different pieces, or spheres, by its social, political and geopolitical contradictions, so that the overall logic of reality, if any, is split into different (sub)logics which have to find a way to interact, not looking for the ghost of a claimed but impossible self-founding unity, but by making the sub-domains they represents negotiate.

The extent to which these (sub)logics succeed in cooperating, or, instead they deepen their conflicts is a *measure of crisis*. The ideology of *globalization* itself is the main witness to the fact that capitalism claims to have a universal logic. On the contrary, as we shall see, globalization is the extensional effect of a predominant specific logic, that is, of predominant specific intensional dynamics, namely that of imperialism, which in turn generates many local competing replicants.

I am aware that I am speaking of analogies or even suggestions. Anyway, scientific discovery goes on by means of analogies and suggestions.

3 Global behaviour and local behaviour

A similar subdivision may appear within a single (sub)logic, so that one can identify a *local logical behaviour* within a global one.



In Logic, a striking example of the above situation is given by an interesting approach which, to some extent, is an intermediate one between Labelled Deductive Systems and Linear Logic. It was proposed by P. Miglioli who introduced a modal operator $\mathbf{T}$ upon the axioms of *Constructive Logic with Strong Negation*, CLSN (see [21]), such that $\mathbf{T}(\alpha)$ means "α is classically valid". The resulting logic was called E_0. Therefore, the operator $\mathbf{T}$ carves out a classical local world within a non-classical one[2]. The original aim of the operator $\mathbf{T}$ was to distinguish two specific cognitive modalities towards. A first modality relates to the assumption of cognitive data. When objects are considered from this point of view, it is not relevant the fact that they are possibly the result of a cognitive subject's constitutive activity. On the contrary, the second cognitive modality is adopted whenever one is interested in analyzing or synthesizing the object of knowledge, so that their modes of constitution are to be taken into account. Therefore, these two different cognitive modalities define *two different logical behaviours of the objects of knowledge*[3].

Given these motivations, it was surprising to find exactly the same mechanisms in a completely different field. I mean Rough Set Theory (see [27] and [26]), which is the main interest shared by Mihir Chakraborty and me. Indeed, this theory provides a good example of a global logics embedding a local logic which describes the behaviour of a specific sub-domain of discourse.

I briefly introduce rough sets in a general manner.

Given a universe U of objects, a set P of properties and a binary relation $R \subseteq U \times P$, let us call the structure $\mathbf{O} = \langle U, P, R \rangle$ an *observation system.* Intuitively, if xRp, for $x \in U$ and $p \in P$, then property p is manifested by x (or, can be observed of x). We shall deal with observation systems later on. By now, let us notice that one can transform an observation system into a relational system between elements of U, i. e., $\mathbf{R} = \langle U, R \rangle$ with $R \subseteq U \times U$, which is a particular observation system where the behaviour of objects with respect to other objects is observed (see [26]). Intuitively, if xRy, for $x, y \in U$, then the object y is observed, or perceived, or manifested together the object x. Or, which is the same, y lays in the *perception field* of x. For this reason, let us call $\mathbf{R}$ a *perception system.*

For any $X \subseteq U$, the set $R(X) = \{y : \exists x \in X(xRy)\}$ is called the *R-neighborhood* of X. If $X = \{x\}$ we shall write $R(x)$ instead of $R(\{x\})$.

In Rough Set Theory, any subset X of a universe U is perceived not directly but as the result of the application of two operators, the *lower approximation*

[2] Cf. [18]. P. Miglioli was not interested in de-structuring/re-structuring Classical Logic or Intuitionistic Logic, but the syntactic side of his works was characterised by imposing classical behaviours to specific parts of a logical system (as in E_0 and its supersystems) or by adding superintuitionistic principles to constructive systems. I mean principles such as the the Kreisel-Putnam principle or the Kuroda formula (see below).

[3] Think of the specification of a program with output A, or an axiom which says that a certain mathematical object exists, for instance the relative pseudocomplementation of any two elements of a lattice. Of course we can use those cognitive objects to reason about output A (for instance as an input for another program to be specified) or to prove theorems by assuming the *platonistic* existence of the relative pseudocomplementation. These are cognitive activities far different from the construction of an algorithm to output A or to compute the relative pseudocomplementation.

$(lR)(X) = \{y : R(y) \subseteq X\}$ and the *upper approximation* $(uR)(X) = \{y : R(y) \cap X \neq \emptyset\}$. We say that two subsets A and B of U are *rough equivalent*, in symbols $A \equiv_r B$, if $(lR)(A) = (lR)(B)$ and $(uR)(A) = (uR)(B)$. A *rough set* is an equivalence class modulo $\equiv_r$.

Particularly in classical Rough Set Theory, R is an equivalence relation with the following meaning; if xRy then the objects x and y are indiscernibile by means of the properties that they manifest to us. Therefore, intuitively $(lR)(X)$ and $(uR)(X)$ are patterns, over a domain of discourse, defined by selected properties. From this point of view they are *conceptual patterns*. Besides equivalence relations, a number of other binary relations have been considered. However, it must be noticed that in order that (lR) may be a lower approximation and (uR) an upper approximation, R must be reflexive. From now on reflexivity is understood[4].

In any case, R is to be intended as it was explained above: if xRy then y is always perceived along with x, through the observed properties provided by the perception system $\mathbf{R}$. It follows that if $R(x) = \{x\}$, then x has unique observed properties. Otherwise stated, x can be singled out within the domain.

It is obvious then one deals with rough sets, globally the following three-valued characteristic function is to be applied, for $A \subseteq U$:

$$\chi_A(x) = \begin{cases} 0 \text{ if } x \in -(uR)(A) \\ \delta \text{ if } x \in (uR)(A) \cap -(lR)(A) \\ 1 \text{ if } x \in (lR)(A) \end{cases}$$

However, on the subuniverse $S = \{x : R(x) = \{x\}\}$ the characteristic function takes only values 1 and 0. Indeed, $R(x) \cap A \neq \emptyset$ if and only if $x \in A$, so that $R(x) \subseteq A$. It follows that if $x \in (uR)(A)$ then $x \in (lR)(A)$. The opposite direction is always valid, since $(lR)(X) \subseteq (uR)(X)$. Therefore the *boundary case* $x \in (uR)(A) \cap -(lR)(A)$ cannot happen. In topology, this fact is expressed by saying that isolated points cannot belong to the boundary of any subset[5].

Therefore, the characteristic function is classic on the subuniverse S. Otherwise stated, on S we observe a local classical behaviour embedded in a three-valued environment.

Further, one can show that the definition of the algebraic models for CLSN and the definition of the algebra of the system of all rough sets from a perception system $\mathbf{R}$ use the same filtration clauses (to filter the classical subuniverse) (see [26]), so that if R is a preorder such that for all $x \in U$ there is $y \geq x$ such that $R(y) = \{y\}$, then the resulting rough set system is an algebraic model for E_0 (see [11]).

[4] It is well-known that (lR) is a necessity operator, while (uR) is a possibility operators. Thus as the properties of R vary, one obtains different modal operators, according to well-known results in the relational theory of modal logic. Moreover, (lR) can be considered a topological or pre-topological interior operator, under certain circumstances, and (uR) a closure operator - see below.

[5] Indeed, $(uR)(A) \cap -(lR)(A)$ is the topological boundary of A, whenever (uR) and (lR) are (pre)topological closure and, respectively, interior operators.

Alexander Grothendiek introduced a particular kind of topology to deal with the notion of "it is locally valid that ...". The closure operators $\mathbb{C}$ induced by a Grothendieck topology distributes over meets, not over unions as in usual topologies[6]. From these topologies the logico-algebraic notion of a *Lawvere-Tierney operator* arose. All the above constructions can be made in terms of Grothendieck topologies or Lawvere-Tierney operators (see [26]).

A notable example of a Lawvere-Tierney operator, induced by a particular Grothendieck topology, is the "local operator" $\langle l \rangle$. One obtains that a property P is locally valid at state s in an Intuitionistic Kripke model Int, $s \models_{Int} \langle l \rangle(P)$, if s forces the double negation of P, that is, if $s \in [\![\neg\neg P]\!]$. From the Gödel-Glivenko theorem, one has immediately that any classical tautology is locally valid in any intuitionistic model.

Thus, this example shows that one has to deal with different specific logics and, further, subdivide them between globally and locally valid logics. Furthermore, it gives rise to the question as to the validity of the hypothesis that Classical Logic is the "ultimate" logic.

In fact, we know, from the Kreisel-Putnam principle and the theorem of Kirk (see [14] and [13]), that in between Intuitionistic Logic and Classical Logic there is a ramified hierarchy of intermediate constructive logics, which have been proved to be $2^{\aleph_0}$ with $2^{\aleph_0}$ maximal elements (cf. [19, 20]). Then the question arises as if this transfinite hierarchy is bound to converge to Classical Logic. The answer is: "Yes, in the propositional case. No in the predicative case".

Indeed, Intuitionism is coherent with the negation of the formula (*) $\forall x(\alpha(x) \vee \neg\alpha(x))$ which is a classical theorem. It follows that the Gödel-Glivenko Theorem is not valid in the predicative case. Therefore, one can extend Intuitionism towards an anti-classical direction[7]. It follows that the problem of a ultimate logic finds a serious problem even at a formal level.

On the other way around, we can have a reversal of situation. For instance, it is well-known that in Quantum Mechanics some phenomena can be globally but not locally valid (see, for instance, [10]). The failure of distributivity in Quantum Logic is due to this fact. In [2] an intuitive semantics for Quantum Logic is introduced by means of "proximity spaces". A proximity space is nothing but a perception system $\langle U, R \rangle$ such that R is a similarity relation (i. e. reflexive and symmetric). In the cited paper, J. B. Bell argues that the failure of the distributivity law in Quantum Logic is a consequence of the way in which phenomena are perceived. In this approach phenomena, or "events" in Bell's terminology, are *quanta at a location*. Let $p \in U$, then a quantum at location $p \in U$ is p together with all the other points that are similar to p. Thus a quantum at location p is the R-neigbourhood $R(p)$. A quantum is intended as the "minimum

[6] Intuitively, if a property is locally valid on some sets A and B, than it is locally valid on their subset $A \cap B$ (i. e. $\mathbb{C}(A \cap B) = \mathbb{C}(A) \cap \mathbb{C}(B)$), while it could be invalid on their superset $A \cup B$.

[7] To avoid this situation one must add to Intuitionism the Kuroda formula $\forall x \neg\neg\alpha(x) \longleftrightarrow \neg\neg\forall x \alpha(x)$, all α.

perceptibilium" at a certain location. Unions of quanta give rise to the space of events ("assemblage of quanta") $Q(\mathbf{R})$, which is an ortholattice.

In a quantum assemblage $Q(\mathbf{R})$, the following "localization property" fails:

- if a and b cover U that is, if $a \vee b = U$, then for each element z of $Q(\mathbf{R})$, the set $\{a, b\}$ "localizes" to a cover $\{a \wedge z, b \wedge z\}$.

Otherwise stated, the distributivity law $(a \wedge z) \vee (b \wedge z) = (a \vee b) \wedge z$ fails in a quantum assemblage. It follows that the Persistence Property which states that phenomena are valid also in subspaces, fails too. Therefore, this semantic is able to model the fact that in Quantum Mechanics we can have properties that hold globally, but not locally. Indeed, we can have properties P_1 and P_2 such that the disjunction $P_1 \vee P_2$ holds in the whole space of events, but does not hold in sub-parts of the space. Actually, if S is such a sub-part then we may have a superposition of P_1 and P_2 while S cannot be split into two sub-parts, one fulfilling P_1 and the other P_2[8].

If we go on social phenomena, we can see, for instance, that neo-liberal ideologists maintain, on the contrary, that common sense logic applies to States and governments. By way of example, sovereign debts and common people's debts would share the same logic. This is a complete nonsense, since a government can mint money (in some cases freely, think of the USA), while common people cannot. Indeed, a number of characteristics of a government's logic do not apply to common people and vice-versa. We have seen that this happens in systems of formal logic too, that is, in systems where contradictions cannot even exists[9]. All the more so, it happens in more complex and intrinsically contradictory domains, such as capitalistic societies.

4 Dynamics

It must be noticed that the Kripke models which make the above classical principle (*) fail, have increasing domains associated to increasing states (i. e., if $s \lneqq t$ then $D(s) \subsetneqq D(t)$, where $D(x)$ is the set of individuals associated with cognitive state, or possible world, x). From this situation a sort of diagonal argument follows. Indeed, in the presence of variability, not only of the cognitive states (represented by the possible worlds in Kripke models) but also of the very objects of the universe of discourse (the domains associated to the cognitive states), any ultimate logic, as Classical Logic is supposed to be, seems to disappear from the horizon.

More in general, variability, hence dynamics, is strongly connected to the possible failure of "limit" principles, such, for instance, idempotence. But limit

[8] In other terms, we cannot have S' and S'' such that $S' \cup S'' = S$ and $S' \models P_1$, $S' \models P_2$, which is a condition explicitly required by the Kripke-Joyal semantic in order for S to force $P_1 \vee P_2$. In this case $U \models P_1 \vee P_2$ but $P_1 \vee P_2$ does not localise to all the sub-parts of U.

[9] Formal systems arise to overcome contradictions, think of Russell's Ramified Type Theory, or ZF.

principles lay at the heart of mathematics, so that some scholar maintains that there cannot exist a "mathematics of transientness". Again, we face a contradiction between pragmatics and formal means.

Dynamics has been a long discussed problem by professor Chakraborty and me, because we are strongly interested to model the formation of conceptual patterns not only from the point of view of a single cognitive subject, but also by means of the interaction between multiple cognitive subjects. We called this approach *dialogical*.

This interest developed within our studies in Rough Set Theory. However, in Rough Set Theory, usually only one cognitive subject is assumed.

So far, therefore, we have seen that a total cognitive subject is fragmented into many single cognitive subjects who, in turn, may have different cognitive attitudes: eventually a global and a local attitude. Now we face the problem of how these subjects can interact.

By now, we start with a more complex, but still simple, situations. Namely, we assume that there can be different points of view at the same point in time, or a single point of view can give different observations in different points in time. Further, we can think also of different points of view over different points in time.

In order to model these situations we have to consider *Dynamic Spaces*, $\langle \mathfrak{U}, \mathfrak{R} \rangle$, where I is an initial segment of $\mathbb{N}$, say $I = \{1, ..., n\}$, $\mathfrak{U} = \{U_i\}_{i \in I}, \mathfrak{R} = \{R_i\}_{i \in I}$ and for each i, U_i is a set while $R_i \subseteq U_i \times U_i$.

It is obvious that dynamic spaces are a basis to develop a *dialogical approach* to conceptual patterns formation. On such a basis one can develop specific cases (the list is partial).

A) **Approximation of relations**.
Upon dynamic spaces one can define an n-ary Relational Approximation Triple, $\mathbf{RA}(\mathfrak{U}) = \langle \mathfrak{U}, \mathfrak{R}, Z \rangle$, where Z is a relation point-wise defined on $\prod_{i=1}^{n} U_i$ by:

$$\langle \langle x_1, ..., x_n \rangle, \langle y_1, ..., y_n \rangle \rangle \in Z \textit{ iff } \langle x_i, y_i \rangle \in R_i, \textit{ all } i.$$

Thus, for any $R \subseteq \prod_{i=1}^{n} U_i$, we set, as usual:

$$(lZ)(R) = \{\langle x_1, ..., x_n \rangle : Z(\langle x_1, ..., x_n \rangle) \subseteq R\}.$$

Hence, $\langle x_1, ..., x_n \rangle \in (lZ)(R)$ whenever $\langle \langle x_1, ..., x_n \rangle, \langle y_1, ..., y_n \rangle \rangle \in Z$ implies $\langle y_1, ..., y_n \rangle \in R$. Dually for $(uZ)(R)$.
In the case $n = 2$, so that $\mathbf{RA}(\mathfrak{U}) = \langle \{U_1, U_2\}, \{R_1, R_2\} \rangle$, a modal-algebraic characterization of this approximation operator can be given: $(lZ)(R) = (R_1^{\smile} \longrightarrow R) \longleftarrow R_2$, where $^{\smile}$ is the reverse operator, $\longrightarrow$ is the right residuation and $\longleftarrow$ is the left residuation (cf. [22, 23]).

Since a set $A \subseteq U$ can be represented as a particular binary relation, called a *right cylinder*, namely $A \times U$, approximation operators in usual Rough Set Theory happens to be particular cases of those induced by approximation of relations.

B) **Consensus operators**

Let us call a dynamic space where for each $i, j \in I$, $U_i = U_j$, *Uniform Domain Dynamic Space* (UDDS). On such observation systems one can define lower and upper approximations according to a certain number m of observations, for $m \in I$. We call these operators *contraction* and, respectively, *expansion*.

1. (*Contraction*): We say that $x \in \varkappa^m(A)$, for $1 \leq m \leq n$, if every y such that $\langle x, y\rangle \in R_i$ belongs to A, at least in m cases. Otherwise stated: $x \in \varkappa^m(A)$ if $R_i(x) \subseteq A$ for at least m indices.
2. (*Expansion*): We say that $x \in \varepsilon^m(A)$, for $1 \leq m \leq n$, if A contains at least a y such that $\langle x, y\rangle \in R_i$ in at least $n + 1 - m$ cases. Otherwise stated: $x \in \varepsilon^m(A)$ if $R_i(x) \cap A \neq \emptyset$ for at least $n + 1 - m$ indices.

Not much is known about the algebraic and topological properties of these operators. Such properties vary according to those of the relations R_i and to the index m of the operators. However the following was proved (see [24]):

1. If $\mathfrak{R}$ is a family of reflexive relations then $\varkappa^1$ is a pretopological *non idempotent* interior operator *which does not distribute over meets*. That is, it enjoys normality, $\varkappa^1(U) = U$, co-normality, $\varkappa^1(\emptyset) = \emptyset$, isotonicity, $A \subseteq B$ implies $\varkappa^1(A) \subseteq \varkappa^1(B)$, and it is deflationary, $\varkappa^1(A) \subseteq A$.
2. If $\mathfrak{R}$ is a family of preorders then $\varkappa^1$ is a pretopological *idempotent* interior operator which does not distribute over meets.
3. if $\mathfrak{R} = \{R\}$, and R is reflexive, then $\varkappa^1$ is a pretopological *non idempotent* interior operator.
4. if $\mathfrak{R} = \{R\}$, and R is a preorder, then $\varkappa^1$ is a topological interior operator.

The latter result is well-known and in this context it is a straightforward conclusion from the previous ones.

C) **Multiple-source Approximation Systems**

From a UDDS $\langle U, \mathfrak{R}\rangle$ one can define four basic approximation operators, two weak (labelled by "w") and two strong (labelled by "s"):

- $(sl\mathfrak{R})(X) = \bigcap_{i \in I}(lR_i)(X)$; $(wl\mathfrak{R})(X) = \bigcup_{i \in I}(lR_i)(X)$;
- $(su\mathfrak{R})(X) = \bigcap_{i \in I}(uR_i)(X)$; $(wu\mathfrak{R})(X) = \bigcup_{i \in I}(uR_i)(X)$.

The resulting system is called a *Multiple-source Approximation System* (MSAS).

In various papers, Md. Aquil Khan and Mohua Banerjee developed a syntax and a semantic for a logic of MSASs (see for instance [12]). Anyway, by definition, one obtains that given a MSAS such that $I = \{1, ..., n\}$, $(sl\mathfrak{R})(X) = \varkappa^n(X)$.

We stop here our brief survey. Just a few comments. First of all, at present we are far from a real dialogical system, because such a system should provide a strategy to dynamically change the different points of view, according to competing observations and a given goal. Anyway, some seeds have been planted that can be used for future enhancements.

Moreover, we can observe that, in the general case, when we deal with operators induced by dynamic spaces, idempotence can fail. To me, this is a "natural consequence" of dynamics and it is worth discussing it.

In usual topology, concepts like *closure* and *interior* are idempotent (and, moreover, they distribute over unions and meets, respectively). The closure of a set X, $\mathbb{C}(X)$, joins to X its topological *boundary* (intuitively, the set of points which are neither inside nor outside X, but are "close" to some point in X). In usual topology, this operation does not produce a new boundary, so that $\mathbb{C}(\mathbb{C}(X)) = \mathbb{C}(X)$, that is, the closure of a closed set does not add new information. Otherwise stated, $\mathbb{C}(X)$ is a *least fixedpoint* of the operator $\mathbb{C}$. A least fixedpoint is a limit point beyond which new application of the operations which have reached it, do not produce anything new. The familiar operations of addition and multiplication, themselves, are *limit operations* (technically they are co-limits and limits), as like as the corresponding operations of meet and union, or the logical $\wedge$ and $\vee$.

On the contrary, in a dynamic situation, the closure of a set could generate a new boundary, and so on: $X \subset \mathbb{C}(X) \subset \mathbb{C}(\mathbb{C}(X))...$. But here we enter into a much less explored mathematical field: that of *pre-topologies*. This field is not new (see [3]) but is now receiving a renewed interest where dynamics plays a main role (see, for instance, [29, 30]).

Prior to those studies, pretopological operators have been considered in Modal Logic, basically because of two situations. First of all, some logical modalities must be modeled by Kripke models where the accessibility relation R is not transitive[10]. Transitivity, in fact, induces limit operations because it is a "stabilization property": everything which is directly or indirectly connected with a set X is gathered in one shot by $R(X)$, leaving nothing outside.

Secondly, there are interesting modalities which must be modeled by *neighbourhood systems*, that is, by systems of families of subsets of U associated to each element $p \in U$. If $\mathcal{N}_p$ is such a family, it is called the *neighbourhood family* of p. A neighbourhood $N \in \mathcal{N}_p$ gathers the elements (called "neighbours") that are close to p under some "vicinity" conditions, that is, from a particular point of view (see [4]). These neighbourhood systems induce topological operators just if some "stabilization conditions" are fulfilled. Namely, on top reflexivity ($p \in N$ for all $N \in \mathcal{N}_p$), any neighbourhood $\mathcal{N}_p$ must have the structure of a filter, that is, (i) if $N \in \mathcal{N}_p$ and $N \subseteq N'$ then $N' \in \mathcal{N}_p$, and (ii) if $N, N' \in \mathcal{N}_p$, then $N \cap N' \in \mathcal{N}_p$. It is not difficult to sense transitivity in condition (i). Condition (ii), instead, is linked to distributivity, hence to isotonicity. But one can easily verify that (i) and (ii) may fail in dynamic situations (see [26, **?**]). Particularly, the operators $\varkappa^m$ and ε^m are induced by neighbourhood systems whose neighbourhood families are structured as filters just in particular cases.

What does this story suggest to social analysis?

[10] A Kripke model is essentially a perception system $\langle U, R \rangle$ equipped with an interpretation function.

Marx's scientific approach follows the methods of the exact sciences developed in the same historical period. Actually, those approaches operate through limit cases, which are very useful to abstraction and unification and, therefore, to the modeling of reality, but that are hardly able to deal with dynamic and transient phenomena. Phenomena which are inherently characterized by contradictions.

To me, Marx was aware of this triangular trade-off between informal rigour, completeness and coherence, that is the fact that a theoretical description of the capitalistic society, which is intrinsically a conflicting, dynamic and contradictory reality, can be complete only if it is not coherent, and vice-versa. Indeed, Marx's work is not systematic and, to me, it is not "systematizable". As a matter of fact, it was Kautsky who arranged Marx's thought into a system, what we call "Marxism", while Marx himself abandoned the idea of publishing, over sixteen years, the second, the third and the fourth volumes of "Capital". Indeed, after publishing the first volume, capitalism entered into a systemic crisis, called Great Depression (1873-1895), a systemic crisis which dramatically changed the capitalistic world and modes of operation and which continued with World War I, the 1929 crisis, World War II and, eventually, finished with the hegemony of the USA over the capitalistic world. Such hegemony is now passing through a systemic crisis, since August 1971 when Mr Nixon announced the incovertibility of the dollar to gold.

Indeed, crises are much longer than developing, in capitalism. Therefore, although the theoretical means provided by Marx's and Engels' analyses are useful and necessary to be guided in a complex reality without going blindly and ending up in chaos, nonetheless we are continuously required, by the political praxis, to remember that reality is much more complex and cannot be reduced to its models. This was Lenin's exhortation to the other Marxist revolutionaries. For instance, he recognized the elegance of Bucharin's or Rosa Luxemburg's programs, from a formal and doctrinal point of view, but he severely criticised their inadequacy with respect to the capitalistic and imperialist reality, which was described by Lenin as "shapeless, unpleasant and not harmonious"; in other terms hardly consistent with the symmetries of the scientific concepts used to analyse it. It follows that concrete analysis orders us to look at the reality in all its complexity and contradictions by mining into the several layers of the "symbolic orders" which let us find our way at different levels of abstraction.

To sum up, we can use *limit* concepts only if we are aware that they describe a reality which is, on the contrary, continuously dismembered and that the stable situations described by those limit concepts are either hypothetical or short-term.

5 Contradictions

So far, we have seen that the idea of an ultimate logic faces a number of hard problems. Not only reality is interpretable from different points of view, but it itself is fragmented into specific sub-domains, each one with its own specific logic which, in turn, may embed local logics. To rule reality, therefore, means to try to reassemble the fragments.

We have seen how this can be done by making different formal logics interact with each other. However, dynamics makes interaction a hard job, in that dynamics requires new instruments which mai fail to enjoy the usual *nice properties* of "limit operations" which describe "stable" situations. We have seen the case of idempotency and isotonicity is bound to fail in dynamic situations, as well (see [26])[11].

However, logical pluralism is a "democratic" approach to the problem, because it recognizes that every logic has its own legitimacy. Consider, instead, capitalism. Although it is an inherently conflicting social relations, and although conflicts induce a strong dynamics which, in turn, continuously induces fractures, capitalism cannot recognize its contradictions, even if it continuously tries to bypass them. Otherwise stated, capitalism has to act as a meta-system which tries to master and manage the fragmenting forces that are generated by its systemic contradictions. Therefore, in a literal sense, capitalism is a totalitarian system, a *unique thought*.

But the problem for capitalism, is that it is *rooted* into a conflicting logic. Let us start with Marx:

> "*Direct force, outside economic conditions, is of course still used, but only exceptionally. In the ordinary run of things, the labourer can be left to the natural laws of production, i.e., to his dependence on capital, a dependence springing from, and guaranteed in perpetuity by, the conditions of production themselves. It is otherwise during the historic genesis of capitalist production. The bourgeoisie, at its rise, wants and uses the power of the state ...* [12].

However, now we know that the "ordinary run of things" does not last for long, while crises are longer and longer and more and more frequent. It follows that the underlying double nature of capitalism is unveiled more and more frequently and for longer and longer periods of time.

What is this "double nature"? We understand it from the above fundamental passage: what we call "Capitalism" is the result of the interaction between the *capitalist power* and the *territorial power*. Indeed, as Giovanni Arrighi and David Harvey point out (with some relevant differences), capitalism is the result of a *political exchange* between the Power of Money and the Power of Territory (see [1, 7]). These two powers need each others, but they have different logics: the capitalist logic is focused on accumulation whereas the territorial logic deals with state formation, politics, culture, direct force. The former is a logic of *spaces-of-flows* while the latter is a logic of *spaces-of-places*.

[11] We could be required to renounce other nice properties, if a local logic applies to particular elements of the universe of discourse. For instance, if we accept that atomic formulae cannot have a constructive proof - because they do not have a logical structure - then for any atomic formula p we should have $p \longleftrightarrow \mathbf{T}(p)$, so that Uniform Substitution cannot be applied in the presence of the logical constant $\vee$.

[12] *Capital*. Vol I, Part VIII. Primitive accumulation.

This necessity to join two different logics generate conflicts, which must be mastered and managed by means of that very impossible unity which springs them.

In capitalism there are two types of conflicts: vertical, i. e., class struggle at large, and horizontal, i. e., inter-capitalist conflicts. Indeed, Lenin and Mao were successful because they were able to wedge political class struggle between the inter-capitalist conflicts.

Capitalistic conflicts are conflicts for power:

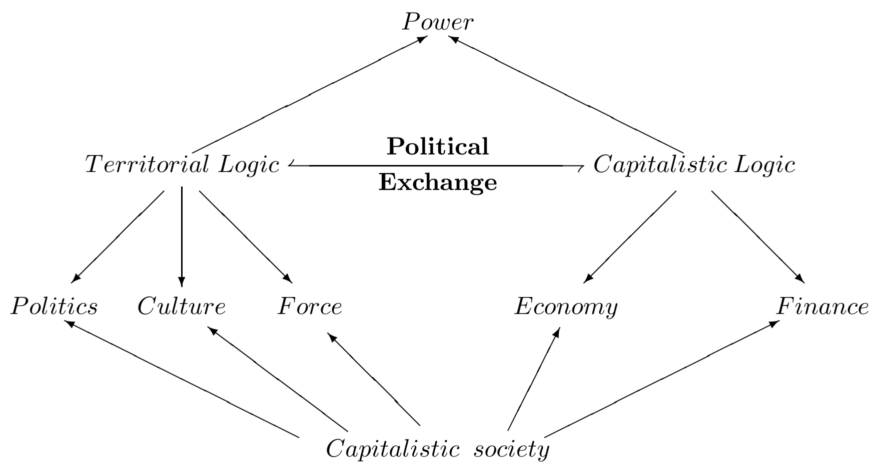
> "*To accumulate, is to conquer the world of social wealth, to increase the mass of human beings exploited by him, and thus to extend both the direct and the indirect sway of the capitalist.* Footnote*: Taking the usurer, that old-fashioned but ever renewed specimen of the capitalist for his text, Luther shows very aptly that the love of power is an element in the desire to get rich*"[13].

Therefore, both capitalist logic, **M**, and territorial logic, **T**, are aimed at *power*. What we, *ex-post*, call "power strategies" is the application, often by trials and errors, of these logics. The fact that **M** and **T** follow different logics, use different means but have the same objective and must take into account each other, is a main source of contradiction which undermine their alliance, continually. The political power is performed within a *territorial fix*, while the capitalist power needs *externalities*. This contradiction breaks the capitalist totality into sub-domains with sub-logics. Economy and finance for **M**, politics, culture, force for **T**. Each sub-domain experience its own conflicts which are managed by domain-specific means and strategies, so that one can observe both inter-domain and intra-domain conflicts (and alliances).

Giovanni Arrighi shows a number of examples of the weaving and parting of these logics. Moreover, he proves that the success of one logic may advantage the other's target, and the other way around. A striking example is given by the conquest of India prompted by Great Britain's territorial logic, which made it possible the development of industrial capitalism in Europe and the USA, under the British hegemony, while the contemporary capitalist logic of France prevented the government from accepting the plan for conquering India, proposed by governor Joseph-Franois Dupleix, rejected by strict cost-benefit calculi. Capitalistic logic did not allowed French decision makers to foresee that the richness robed by the British after Plassey would permit the rising of Great Britain as the pre-dominant industrial, financial, political and military power in the world.

To sum up, the general schema of capitalistic dynamics is something like this:

[13] *Capital*, Vol. I, Chapter Twenty-Four: Conversion of Surplus-Value into Capital.



This pattern is repeated by different nation-states, and this leads to inter-imperialistic conflicts, crises and chaos until a particular power becomes hegemonic.

Marx described the sequence of hegemonic powers in Western capitalism: the Italian city-states, Venice, Spain, Holland, Great Britain, up to the then pretender USA. Giovanni Arrighi, following Fernand Braudel, observed that in between each hegemonic cycle there was a long period of systemic crisis due to a major divergence between **M** and **T**. Thus, a question arises as how to describe the dialectic relationships between the two powers. My suggestion is to borrow the notion of an "*adjunction*" from Mathematics.

We meet adjunctions in a natural way when we deal with an observation system $\langle U, P, R\rangle$. In fact consider the following *basic constructors* (see [26]).

- $\langle e\rangle : \wp(P) \longmapsto \wp(U); \langle e\rangle(Y) = \{g \in U : \exists m(m \in Y \;\&\; g \in R^{\smile}(m))\}$;
- $[e] : \wp(P) \longmapsto \wp(U); [e](Y) = \{g \in U : \forall m(g \in R^{\smile}(m) \Longrightarrow m \in Y)\}$;
- $\langle i\rangle : \wp(U) \longmapsto \wp(P); \langle i\rangle(X) = \{m \in P : \exists g(g \in X \;\&\; m \in R(g))\}$
- $[i] : \wp(U) \longmapsto \wp(P); [i](X) = \{m \in P : \forall g(m \in R(g) \Longrightarrow g \in X)\}$;

We call $\langle e\rangle$ and $[e]$, *extensional* Possibility and Necessity, respectively, while the constructors labeled by "i" are their *intensional* counterparts. The following fundamental relationships hold:
(1) $\langle i\rangle(Y) \subseteq X$ iff $Y \subseteq [e](X)$; (2) $\langle e\rangle(X) \subseteq Y$ iff $X \subseteq [i](Y)$; (3) $Y \subseteq [[e]](X)$ iff $X \subseteq [[i]](Y)$.

These properties tell us that if we set $\mathbf{U} = \langle \wp(U), \subseteq_U\rangle$ and $\mathbf{P} = \langle \wp(P), \subseteq_P\rangle$, then:

- $\langle i\rangle$ and $[e]$ form a *Galois adjunction* between $\mathbf{U}$ and $\mathbf{P}$: $\mathbf{U} \dashv^{\langle i\rangle,[e]} \mathbf{P}$;
- $\langle e\rangle$ and $[i]$ form a *Galois adjunction* between $\mathbf{P}$ and $\mathbf{U}$: $\mathbf{P} \dashv^{\langle e\rangle,[i]} \mathbf{U}$.

In more general terms:

- $\langle i\rangle$ is *left adjoint* to $[e]$ which, in turn is *right adjoint* to $\langle i\rangle$, $\langle i\rangle \dashv [e]$;

– $\langle e \rangle$ is left adjoint to $[i]$ which, in turn is right adjoint to $\langle e \rangle$, $\langle e \rangle \dashv [i]$.

The subset relation $\subseteq$ can be seen as a *transformation* of the elements of the powersets $\wp(P)$ and $\wp(U)$. Thus a Galois adjunction says what happens to the structure (namely a *category*) $\mathbf{P}$ when a transformation takes place in the structure $\mathbf{U}$, and vice-versa[14]. Functors make it possible to move from a structure (a category) to another structure. This is why they are important to deal with competing or/and collaborating structures. Particularly, adjoint functors guarantee that transformations in one structure are reflected by transformations in the other structure[15]. The general adjunction schema between two categories $\mathbf{C}$ and $\mathbf{D}$ is the following:

A pair of functors $F : \mathbf{D} \to \mathbf{C}$ and $G : \mathbf{C} \to \mathbf{D}$, and a family of bijections $\mathcal{H}om_{\mathbf{C}}(FY, X) \cong \mathcal{H}om_{\mathbf{D}}(Y, GX)$ which is natural in the variables X and Y.

Without entering into technical details, it is worth noticing that the clauses of bijection and naturality say that the effect of F on Y is "mirrored without discontinuity" by that of G on X. One can go so far as to say that adjoint functors establish a *dialectic* between structures.

As a matter of fact, one of the mathematicians who presented adjoint functors, William Lawvere, maintained that adjoint functors are able to express materialistic dialectic. It is worth noticing that this claim was supported by strong mathematical arguments (see [15, 16]).

I do not know to what extent this dialectic is "materialistic"[16]. All the same, I maintain that capitalism is characterised by a continuous attempt to establish an adjunction between $\mathbf{T}$ and $\mathbf{M}$, and keep it, because transformations in the political and social structure should induce corresponding transformations in the

[14] A *category* is a class of objects equipped with a class of morphisms, i. e., *transformations* between objects, fulfilling a few properties (existence of the identity morphism for any object, composition of morphisms according to their domains and codomains, associativity of the composition). The class of all morphisms from object A to object B in a category $\mathbf{C}$ is denoted by $\mathcal{H}om_{\mathbf{C}}(A, B)$.

[15] A functor maps objects and morphisms of one category onto objects and morphisms of the other, preserving identitities and composition - for instance, $\langle i \rangle$ is a functor between $\mathbf{U}$ and $\mathbf{P}$, while $[e]$ is a functor between $\mathbf{P}$ and $\mathbf{U}$; $\langle i \rangle$ maps elements of $\wp(U)$ to elements of $\wp(P)$ (the other way around for $[e]$) and $\subseteq_U$ to $\subseteq_P$ (and vice-versa for $[e]$).

[16] When, in the Seventies, I was attending Lawvere's lectures on adjoint functors and dialectical materialism, I was perplexed about the possibility to use formal concepts to interpret Hegel's and Marx's dialectic. After much thought on the current systemic crisis, I recognized that Lawvere's suggestion was absolutely valuable. An interesting approach, inspired to Lawere's intuitions, with several intersections with my present analysis, is given in [8]. However, in that paper it is maintained that there is a global logic in which, particularly, no assumptions are needed because the existence of the universe is not a philosophical assumption itself. I agree on that, but I insist that between the physical and organic world and the social world there is an "ontological jump" induced by *work*, which is, according to Lukács, a "teleological act" (see [17]).

economic and financial structure, in a consistent way, and vice-versa. This is the nature of the alleged *totality* of capitalism.

If we accept that the funding relationship of capitalism, that is, the political exchange **T**-**M**, forms a "dialectic" adjuntion, then we should look for the the following features which are suggested by the way adjoint functors have been interpreted and used in working Mathematics and Logic (we have seen one example above):

Left adjoint - F	Right adjoint - G
extension	intension
ontology	general rules and relations
possibility	necessity
dilemma	choice
past	future
subdivision	integration
sum (juxtaposition) of factors	multiplication (synergy) of factors
closure (action of F after G)	interior (action of G after F)

Thus, any capitalistic phenomena should be analised as an adjoint pair $F \dashv G$. For instance, as $\mathbf{T} \dashv \mathbf{M}$, if it is the territorial dynamics which acts as left adjoint (hence as a premise, an *action*), whereas the economical dynamics acts as right adjoint (as a conclusion, a *retraction*).

In such a case, one obtains, depending on the phenomenon, a pair $\langle$*extension*, *intension*$\rangle$ (material, i. e. physical and geographical, facets, on the left side, and actions for a purpose in that physical environment, on the other side); or $\langle$*possibility*, *necessity*$\rangle$ (for instance, the territorial pre-requisites of an action of valorisation[17], on the one side, and the economic components which give rise to that action, on the other); or $\langle$*integration*, *subdivision*$\rangle$ (for instance, the necessity, on the part of the territorial power, to keep social cohesion, in contrast with the tendency on the part of the capitalist power to atomize). And so on.

On the other way around, we could deal with a pair $\mathbf{M} \dashv \mathbf{T}$, with reversed roles, if it is the economic side which is a basis for a political action, (for instance, valorisation actions can promote territorial expansions). And so forth.

Particularly, the dialectic *past-future* in capitalist accumulation processes is the opposite of what looks to be normal. Indeed, it is the possibility to accumulate tomorrow which guarantees valorisation today. In other terms, the future has feedbacks on the past. Not only a possible, but also a necessary paradox in the capitalist never-ending accumulation process.

Here there is a major depart from the mathematical theory of adjoint functors: usually two mathematical functors cannot be both right and left adjoints to each other.

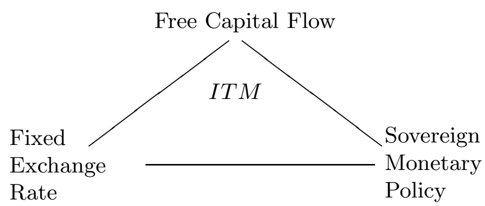
All the same, adjointness describes one part of the story of real capitalistic society. However, the other part is characterised by the breaking of any adjoint

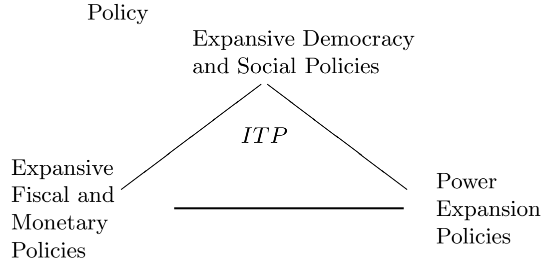
[17] Here understood as the realisation of an increase of value of capital assets, by abuse of language.

relationship, because of the conflicts which are generated by the inherent contradictions of the capitalist social relation.

The reasons of such a breaking are suggested by considering the clause of "naturality" of the families of bijections, provided by the formal definition of adjointness. "Naturality" suggests to check the connections between the transformations which describe the dynamics of **M** and those which describe the dynamics of **T**. Thus we ask: first of all, by what means are these transformations induced and managed in Western capitalism; secondly, how are they applied in the two "structures"; finally, what are the main relationships between these means? Are they "natural", i. e., coherent?

I think that the means can be synthesised in the following list: 1) monetary policies, 2) power politics, 3) democracy and social policies, 4) capital movements. These ingredients may be given an expansive or restrictive dynamics. It is clear that not all the possible combinations are coherent. As a matter of fact we can arrange them into two *inconsistent trinities*. The first, dealing with economical parameters, is the well-known Mundell-Fleming "Unholy trinity", while I have drawn the second upon the first by taking into account political ingredients:





We call the first "Inconsistent Triangle of Money", ITM, and the second "Inconsistent Triangle of Power", ITP.

It is well-known, that Mr Johnson's and Mr Nixon's expansive policies of "guns and butter" led to the so-called *Nixon Shock*, that is, the inconvertibility of the dollar to gold. As a matter of fact, at the same time they tried to avoid the depressive effects of a fiscal and monetary austerity and to increase the expansion of the political and military power of the USA. In other terms, as they were involved in the escalations in Vietnam, the two wormonger Presidents dare not confront the effects of the high social costs of a deflationary policy, to the extent that Mr. Nixon once affirmed: "Now, we are all Keynesian". In that decision the pivotal role was played by the triangular trade-off displayed by ITP. As to the ITM, it was introduced by Robert Mundell e Marcus Fleming in the

Sixties, as a result of the analysis of the very contradictions which led to the Nixon Shock.

Is it obvious that ITP and ITM interact each other in a vicious circle of increasing contradictions. And it is a central problem of the real capitalistic societies to contrast these interlocked inconsistencies. However, this is allowed only to the pre-dominant powers which can discharge offload their contradictions on the subordinate powers.

When one or more of these impossibilities are suspended, one usually says that "the elephants can fly":

> *It is now clear to me that we have followed the dot-com bubble with the 9/11 bubble. Both bubbles made us stupid. [...] The first ended in tears, and so will the second. Because, as the dot-com bubble proved, elephants can fly - "provided it is not very long"*[18].

Although the military Keneysianism avoided a wild "butchery of capitals" (as it would be called by Marx), nevertheless a hard struggle between the Power of Territory and the private financial capital began after the Nixon Shock and lasted until the Reaganomics. In that period the adjunction between **M** and **T** was seriously damaged. However, Mr Reagan gave up social development and with the deregulation a new adjunction relation was established: financialization started tacking flight.

There is no space to account for the stop-and-go policies connected to the effects of the weaving of the two inconsistency triangles[19]. Further, one should split **M** into a financial part and an industrial part, thus making things more complicated[20].

Anyway, financialisation and globalisation are products of the contradiction which are generated by that adjoint relation. The former is a way to increase capital value in spite of its sovraccumulation, by generating *fictitious capital*. The latter has been described as "the biggest heist in history" ([9]).

Adjunction properties cannot formally tell us when the above contradictions explode and how they can be managed. However, we can understand it by looking at the nature of the capitalistic adjuntion relations itself. Namely, it is the very political exchange between **T** and **M** which provides the solution: the bigger is **T** the less material are the "bronze laws" of Economics and the longer ITP and ITM impossibilities can be suspended. This lets elephants fly beyond any expectation. Therefore, this current flight is supported by a specific political, military and diplomatic super-power: the USA. It is US predominance which makes it possible to discharge ITP and ITM impossibilities on world-wide externalities, which represent a six time larger population.

To sum up, a flying elephant is such in dependence of the global relations of force.

From an abstract economic point of view, the conquest of the 28% of world GDP (i. e., India) by the 1.9% (i. e., Great Britain), was a "flying elephant",

[18] Thomas Friedman, "The 9/11 Bubble". The New York Times, December 2, 2004.
[19] Such as the Plaza Accord and the Reverse Plaza Accord.
[20] additional mathematical schemata must then be used, such as pull-backs.

surely not less than the current Three Trillion Dollar War against Iraq (see [31]). But it happened.

References

1. G. Arrighi, **The Long Twentieth Century: Money, Power and the Origins of Our Times**. Verso, 1994.
2. J. L. Bell, "Orthologic, forcing, and the manifestation of attributes". In C. T. Chong & M. J. Wicks (Eds.), **Southeast Asian Conference on Logic**. Elsevier Science Publisher, North-Holland, 1983, pp. 13-36.
3. E. Cech, **Topological Spaces**. John Wiley and Sons, 1966.
4. B. Chellas, **Modal Logic: An Introduction**. Cambridge University Press, 1980.
5. D. M. Gabbay, **Labelled Deductive Systems**. Oxford University Press, 1997.
6. J-Y. Girard, "On the Unity of Logics". *Annals of Pure and Applied Logic*, 59, North-Holland, 1993, pp. 201-217.
7. D. Harvey, **The new imperialism**. Oxford University Press, 2003.
8. M. Heather & N. Rossiter, "Logical Monism: The Global Identity of Applicable Logic". In G. Sica (Ed.), **Essays on the Foundations of Mathematics and Logic**. Polimetrica International Scientific Publisher, 2005, pp-39-52.
9. M. Hudson, **Super Imperialism - New Edition: The Origin and Fundamentals of U.S. World Dominance**. Pluto Press, 2003.
10. M. Jammer, **The Philosophy of Quantum Mechanics**. Wiley, New York, 1974.
11. J. Järvinen, P. Pagliani & S. Radeleczki, "Atomic information completeness in rough set systems". Extended abstract. RST11, 14-16 September 2011, Milano, Italy.
12. Md. A. Khan & M. Banerjee, "An Algebraic Semantics for the logic of Multiple-source approximation Systems". manuscript.
13. R. E. Kirk, "A result on propositional logics having the disjunction property". *Notre Dame Jour, of Formal Logic*, 23, 1982, pp. 71-74.
14. G. Kreisel and H. Putnam, "Eine Unableitbarkeitsbeweismethode fr den in tuitionistischen Aussegenkalkül". *Archiv für mathematische Logik und Grundlagenferschung*, 3, 74–78.
15. F. W. Lawvere, "Adjointness in Foundation". In *Dialectica*, 23 (3/4), 1969, pp. 281-296.
16. F. W. Lawvere, "Towards the Description in a Smooth Topos of the Dynamically Possible Motions and Deformations of a Continuous Body". In *Chaiers Top. Gome. Diff.*, XXXI-4, 1980, pp. 377-392.
17. G. Lukács, **The Ontology of Social Beings**. Merlin Press, 1978.
18. P. Miglioli, U. Moscato, M. Ornaghi and U. Usberti, "A constructivism based on Classical Truth". *Notre Dame Journal of Formal Logic*, 30, 1989, 67–90.
19. P. Miglioli, U. Moscato, M. Ornaghi, M. Quazza and S. Usberti: "Some results on intermediate constructive logics". *Notre Dame Journal of Formal Logic*, 30, 1989, 543–562.
20. P. Miglioli "An infinite class of maximal intermediate propositional logics". *Archive for Mathematical Logic*, 31, 1992, 415-432.
21. D. Nelson, "Constructible Falsity". *Journal of Symbolic Logic*, 14, 1949, 16–26.

22. P. Pagliani, "A modal relation algebra for generalized approximation spaces". In S. Tsumoto, S. Kobayashi, T. Yokomori, H. Tanaka & A. Nakamura (Eds.): **Proceed. Of the 4th International Workshop on Rough Sets, Fuzzy Sets, and Machine Discovery**. November 6-8, 1996, The University of Tokyo, Invited Section "Logic and Algebra", pp. 89-96.
23. P. Pagliani, "Modalizing Relations by means of Relations: a general framework for two basic approaches to Knowledge Discovery in Database". In M. Gevers (Ed.): **Proceed. of the 7th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems**. IPMU '98. Paris, France, July 6-10, 1998. "La Sorbonne", Editions E.D.K, pp. 1175-1182.
24. P. Pagliani, "Pretopology and Dynamic Spaces". In **Proc. of RSFSGRC'03**, Chongqing, R. P. China, 2003. Extended version in *Fundamenta Informaticae*, 59(2-3), 2004, 221-239.
25. P. Pagliani, "Classical principles in non classical frameworks. Logico-philosophical motivations and applications to partial information systems". Invited talk at the International Conference on Logic, Navya-Nyaya and Applications. Calcutta 2000. Now Ch. 9 of [26].
26. P. Pagliani & M. K. Chakraborthy, **A Geometry of Approximation. Rough Set Theory, Logic, Algebra and Topology of Conceptual Patterns**. Springer, 2008.
27. Z. Pawlak, **Rough Sets: A Theoretical Approach to Reasoning about Data**. Kluwer, 1991.
28. A. D. Sokal, J. Bricmont **Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science**. St Martins Pr, 1988.
29. B. M. R. Stadler and P. F. Stadler, *Generalised Topological Spaces in Evolutionary Theory and Combinatorial Chemistry.* Technical Report, Institut für Theoretische Chemie und Molekulare Strukturbiologie, Wien University, 2001.
30. B. M. R Stadler, P. F. Stadler, G. P. Wagner & W. Fontana, "The topology of the possible: Formal spaces underlying patterns of evolutionary change" - www.tbi.univie.ac.at
31. J. Stiglitz & L. Bilmes, **The Three Trillion Dollar War: The True Cost of the Iraq Conflict**. W. W. Norton & Company, 2008.

Piero Pagliani
Research Group on Knowledge and Communication Models,
Via imperia 6, 00161 Roma, Italy.
p.pagliani@agora.it

F-granulation and Generalized Rough Sets: Uncertainty analysis and pattern recognition.

Sankar K. Pal

Abstract. Role of *f*-granulation in machine and human intelligence, and its modeling through rough-fuzzy integration are discussed. Several examples of synergistic integration, e.g., rough-fuzzy case generation, rough-fuzzy clustering, rough-fuzzy classification and rough-fuzzy entropy are explained with their merits and the role of *f*-granulation. Generalized rough sets considering both the set and granules as crisp and/or fuzzy are defined. Grayness and spatial image ambiguity measures using exponential and logarithmic gain functions are stated accordingly. The significance of neighborhood rough sets in dimensionality reduction is explained. The superiority of integration, in terms of performance and computation time, is illustrated for the tasks of case mining in large scale case based reasoning systems, segmenting brain MR images, classifying remotely sensed images and analyzing protein sequences, as examples. The article includes some of the result published elsewhere.

1. Introduction

Rough set theory is a popular mathematical framework for granular computing. The focus of rough set theory is on the ambiguity caused by limited discernibility of objects in the domain of discourse. Granules are formed as objects and are drawn together by the limited discernibility among them. Rough set represents a set in terms of lower and upper approximations. The lower approximation contains granules that completely belong in the set and the upper approximation contains granules that partially or completely belong in the set. Rough set based techniques have been used in the fields of pattern recognition, image processing, data mining and knowledge discovery process from large data sets. Rough sets were found to have extensive application in dimensionality reduction and knowledge encoding particularly when the uncertainty is due to granularity in the domain of discourse. It is also found to be an effective machine learning tool for designing ensemble classifier.

Rough-fuzzy or fuzzy-rough techniques are efficient hybrid methods based on judicious integration of the principles of rough sets and fuzzy sets. While the membership functions of fuzzy sets enable efficient handling of overlapping classes, the concept of lower and upper approximations of rough sets deals with uncertainty, vagueness, and incompleteness in class definition using the notion of granularity.

It may be mentioned that the concept of rough-fuzzy computing has a significant role in modeling the fuzzy-granulation (f-granulation) characteristics of Computational theory of perceptions (CTP) [3,4, 29] which is inspired by the remarkable human capability to perform a wide variety of physical and mental tasks, including recognition, without any measurements and computations. Perceptions are intrinsically imprecise. Their boundaries are fuzzy and the attribute they can take are granules. In other words, perceptions are f-granular.

The present article deals with the significance of rough-fuzzy computing in uncertainty analysis, in general, and pattern recognition and image processing in particular. Different tasks of pattern recognition and machine learning are considered with various real life applications. The organization of the paper is as follows. Section II presents, in brief, f-granulation and the rough-fuzzy approach to granular computation. Section III explains the application of rough-fuzzy granulation in case based reasoning where the problem of case generation is considered. Section IV describes a classification method demonstrating the power of neighborhood rough sets in feature selection from f-granulated space. Section V demonstrates the concept of rough-fuzzy clustering. The problem of segmenting brain MR images is considered, as an example. Section VI demonstrates an application of rough-fuzzy clustering for analyzing protein sequence for determining bio-bases. Section VII deals with generalized rough sets, entropy and image ambiguity measures. Its application to image segmentation, as an example, is shown. Concluding remarks are given in Section VIII.

2. Granular Computation, *f*-granules and Rough-Fuzzy Approach

Rough set theory [5] provides an effective means for analysis of data by synthesizing or constructing approximations (upper and lower) of set concepts from the acquired data. The key notions here are those of “information granule” and “reducts”. Information granule formalizes the concept of finite precision representation of objects in real life situation, and reducts represent the core of an information system (both in terms of objects and features) in a granular universe. Granular computing (GrC) refers to that where computation and operations are performed on information granules (clump of similar objects or points). Therefore, it leads to have both data compression and gain in computation time, and finds wide applications. An important use of rough set theory and granular computing in data mining has been in generating logical rules for classification and association. These logical rules correspond to different important regions of the feature space, which represent data clusters.

In many situations, when a problem involves incomplete, uncertain and vague information, it may be difficult to differentiate distinct elements and one is forced to consider granules. On the other hand, in some situations though detailed information is available, it may be sufficient to use granules in order to have an efficient and practical solution. Depending on the nature of problems and data, granules could be fuzzy rather than crisp; thereby leading to *f*-granulation. Granulation is an important step in the human cognition process. From a more practical point of view, the simplicity derived from granular computing is useful for designing scalable data mining algorithms. There are two aspects of granular computing, one deals with formation, representation and interpretation of granules (algorithmic aspect) while the other deals with utilization of granules for problem solving (semantic aspect). Several approaches for granular computing have been suggested in literature including fuzzy set theory, rough set theory, power algebras and interval analysis. The rough set theoretic approach is based on the principles of set approximation and provides an attractive framework for data mining and knowledge discovery.

For the past few years, rough set theory and granular computation has proven to be another soft computing tool which, in various synergistic combinations with fuzzy logic, artificial neural networks and genetic algorithms, provides a stronger framework to achieve tractability, robustness, low cost solution and close resembles with human like decision making. For example, rough-fuzzy integration can be considered as a way of emulating the basis of *f*-granulation in CTP, where perceptions have fuzzy boundaries and granular attribute values. Similarly, rough-fuzzy entropy can be defined to model various image ambiguities arising from both fuzziness and rough resemblance in gray values and pixels. Since in granular computing computations/operations are performed on granules (clump of similar objects or points), rather than on the individual data points, the computation time is greatly reduced. The results on these investigations, both theory and real life applications, are available [6-8].

Some examples of rough fuzzy computing in clustering, classification, and in defining generalized rough entropy and image ambiguity measures are explained in the following sections together with their characterize features. Some real life applications are also given to demonstrate their effectiveness. Before we describe them, we present briefly the concept of case generation in rough-fuzzy framework as it forms the basic principles of *f*-granulation in several other applications.

3. Rough-Fuzzy Granulation and Case Generation

A case may be defined as a contextualized piece of knowledge representing an evidence that teaches a lesson fundamental to achieving goals of the system. Case based reasoning (CBR) [9] is a novel Artificial Intelligence (AI) problem-solving paradigm, and it involves adaptation of old solutions to meet new demands, explanation of new situations using old instances (called cases), and performance of reasoning from precedence to interpret new problems. It has a significant role to play in today's pattern recognition and data mining applications involving CTP, particularly when the evidence is sparse. The significance of soft computing to CBR problems has been adequately explained by Pal, Dillon and Yeung [10] and Pal and Shiu [11]. In this section we provide an example [12, 13] of using the concept of f-granulation for performing the task of case generation in large scale CBR systems. While case selection deals with selecting informative prototypes from the data, case generation concerns with construction of 'cases' that need not necessarily include any of the given data points.

For generating cases, linguistic representation of patterns is used to obtain a fuzzy granulation of the feature space. Rough set theory is used to generate dependency rules corresponding to informative regions in the granulated feature space. The fuzzy membership functions corresponding to the informative regions are stored as cases.

Figure 1 shows an example of such case generation for a two dimensional data having two classes. The granulated feature space has 32 = 9 granules. These granules of different sizes are characterized by three membership functions along each axis, and have ill-defined (overlapping) boundaries. Two dependency rules: class1 ← L1 ∧ H2 and class2 ← H1 ∧ L2 are obtained using rough set theory. The fuzzy membership functions, marked bold, corresponding to the attributes appearing in the rules for a class are stored as its case.

Unlike the conventional case selection methods, the cases here are cluster granules and not sample points. Also, since all the original features may not be required to express the dependency rules, each case involves a reduced number of relevant features. The methodology is therefore suitable for mining data sets, large both in dimension and size, due to its low time requirement in case generation as well as retrieval.

The aforesaid characteristics are demonstrated in Figure 2 [12, 13] for a real life data set with number of features 10 and number of samples 586012. Its superiority over IB3, IB4 [9] and random case selection algorithms, in terms of classification accuracy (with one nearest neighbor rule), case generation (tgen) and retrieval (tret) times, and average storage requirement (average feature) per case, is evident. The number of cases considered for comparison is 545. Based on the similar concept, Li et al reported a CBR based classification system combining efficient feature reduction and case selection [14]. Note that here the granules considered are class independent (CI). In the next section we describe a classification method where the granules are class dependent (CD).

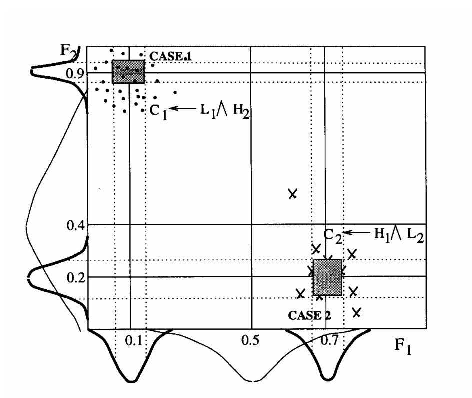


Fig. 1. Rough-fuzzy case generation for a two dimensional data [12]

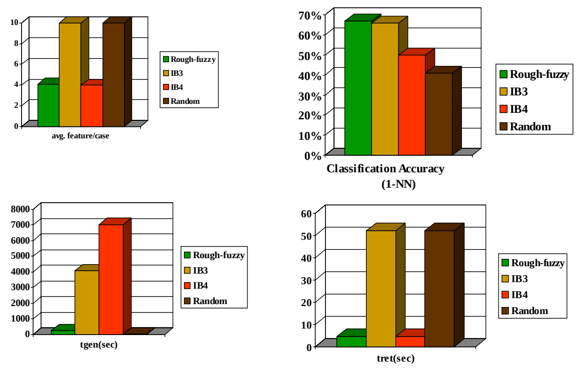


Fig. 2. Performance of different case generation schemes for the forest cover-type GIS data set with 7 classes, 10 features and 586012 samples

4. Rough-Fuzzy Classification

The rough-fuzzy model has three steps of operation as shown in Fig.1.

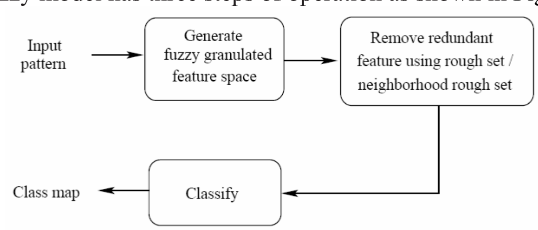


Fig. 3. Schematic diagram for rough-fuzzy pattern classification

The first step generates the class-dependent (CD) fuzzy granulated feature space of input pattern vector. For fuzzy granulation of a feature space containing **L** number of classes, we used **L** number of fuzzy sets (π-type membership function) to characterize the feature values of each pattern vector. Each feature is thus represented by **L** [0,1]-valued membership functions (MFs) representing **L** fuzzy sets or characterizing **L** fuzzy granules along the axis. The π-type MF explores the degree of belonging of a pattern into different classes based on individual features and the granules thus provide an improved class-wise representation of input patterns. The granules preserve the interrelated class information to build an informative granular space which is potentially useful for improved classification for the data sets with overlapping classes. In the granulation process, each feature value is represented with more than one membership values and thus the feature dimension increases. The increased dimension brings great difficulty in solving many tasks of pattern recognition, machine learning and data mining. This motivates for selecting a subset of relevant and non-redundant features. In this regard, we have used the neighborhood rough set [5, 6, 28] (NRS) based feature selection method in the second step (Fig. 3). The advantage in the use of NRS is that it can deal with both numerical and categorical data. NRS does not require any discretisation of numerical data and is suitable for the proposed fuzzy granulation of features. Further, the neighboring concept facilitates to gather the possible local information through neighbor granules that provide better class discrimination information. Thus the combination of these two steps of operations can be a better framework for the classification of patterns in overlapping class environment. The integrated model thus takes the advantage of both class-dependent fuzzy granulation and NRS feature selection methods. After the features are selected, we use a classifier as in the third step of Fig. 3 to classify the input pattern based on the selected features. The aforesaid steps are now discussed in brief.

Let a pattern (object) **F** be represented by n numeric features and be expressed as: $\mathbf{F} = [F_1, F_2, ..., F_n]$. Thus **F** is visualized as a point in n-dimensional vector space. Each feature is described in terms of its fuzzy membership values corresponding to **L** linguistic fuzzy sets. Thus, an n-dimensional pattern vector is expressed as (n x **L**)-dimensional vector and is given by

$$\begin{aligned}\mathbf{F} = \quad & [\mu_1^1(F_1), \mu_2^1(F_1), ..., \mu_c^1(F_1), ..., \mu_{\mathbf{L}}^1(F_1); \\ & \mu_1^2(F_2), \mu_2^2(F_2), ..., \mu_c^2(F_2), ..., \mu_{\mathbf{L}}^2(F_2); \\ & \mu_1^n(F_n), \mu_2^n(F_n), ..., \mu_c^n(F_n), ..., \mu_{\mathbf{L}}^n(F_n)], \quad (c = 1, 2, ..., \mathbf{L})\end{aligned}$$

where $\mu_1^n(F_n)$, $\mu_2^n(F_n)$,..., $\mu_n^n(F_n)$,..., $\mu_L^n(F_n)$ signify the membership values of F_n to **L** number of fuzzy sets along the n^{th} feature axis and $\mu(F_n) \in [0,1]$. That means each feature F_n is expressed separately by **L** number of membership functions In other words, each feature F_n characterizes **L** number of fuzzy granules along each axis and thus comprising $\mathbf{L}^n$ fuzzy granules in an n-dimensional feature space. Fig. 4 shows a crisp visualization of 16 (= 4^2) such class dependent granules using 0.5-cut when the number of classes is four in two-dimensional (F_1 and F_2) feature space. Shape and size of the granules are dependent on the nature of overlapping of classes and class-wise feature distribution.

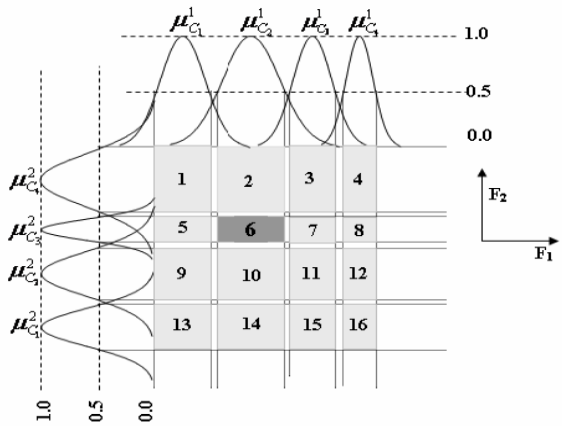


Fig. 4. Sixteen class dependent crisp granules for L = 4 in F1-F2 space

For implementation of the concept of neighborhood rough sets [28] in feature selection, let us assume an information system denoted by I = (U, A) where U (the universal set) is a non-empty and finite set of samples {x1, x2, ..., xn}; A = {C ∪ D}, where A is the finite set of features {a1, a2, ..., am}, C is the set of conditional features and D is the set of decision features. Given an arbitrary xi ∈ U and B C, the neighborhood ΦB(xi) of xi with given Φ, for the feature set B is defined as

$$\phi_B(x_i) = \{x_i | x_j \in U, \Delta^{\mathrm{B}}(x_i, x_j) \le \phi\} \tag{1}$$

where Δ is a distance function.

ΦB(xi) in Eqn. (1) represents the neighborhood information granule centered with sample xi. That is, each sample xi generates granules with a neighborhood relation. For a metric space (U, Δ), the set of neighborhood granules {Φ(xi)| xi ∈ U} forms an elemental granule system, that covers the universal space rather than partitions it as done by Pawlak's rough set (PaRS). A neighborhood granule degrades to an equivalence class when Φ = 0. In this case, samples in the same neighborhood granules are equivalent to each other and neighborhood model degenerates to Pawlak's rough set. Thus NRS (neighborhood rough set) can be viewed as a generalization of PaRS.

Generation of neighborhood depends on both distance function Δ and parameter Φ. The first one determines the shape and second controls the size of neighborhood granule. For example, with Euclidean distance the parameter Φ acts as the radius of the circle region developed by Δ function. Both these factors play important roles in neighborhood rough sets (NRS) and can be considered as to control the granularity of data analysis. The significance of features varies with the granularity levels.

Accordingly, the NRS based algorithm selects different feature subsets with the change of Δ function and Φ value.

Performance of rough-fuzzy feature selection (granular feature space and rough feature selection) is demonstrated with k-NN classifier, as an example, on various data sets. Here we present some results for k = 1 on remotely sensed images where the different regions are highly overlapping and the no. of available training samples is small. Table 1 shows the comparative performance of various models in terms of β value [19] and Davies-Bouldin (DB) value [20] on IRS-1A image and SPOT image with partially labelled samples. Partially labelled means, the classifiers are initially trained with labelled data of six land cover types and then the said trained classifiers are applied on the unlabeled image data to partition into six regions.

Five different models considered are [27]:

- Model 1 : k-NN (with k=1) classifier,
- Model 2 : CI fuzzy granulation + Pawlak's rough set (PaRS) based feature selection + k-NN (with k=1) classifier,
- Model 3 : CI fuzzy granulation + neighborhood rough set (NRS) based feature selection + k-NN (with k=1) classifier,
- Model 4 : CD fuzzy granulation + PaRS based feature selection + k-NN (with k=1) classifier,
- Model 5 : CD fuzzy granulation + NRS based feature selection + k-NN (with k=1) classifier.

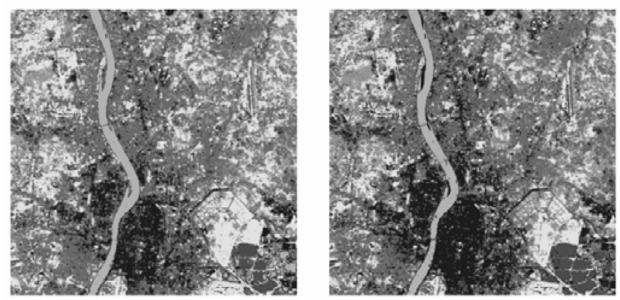
Table 1. Comparative performance of models using k-NN classifier (k=1) with partially labeled data sets (for Φ = 0.45 and Δ = Euclidean distance)

Model	β value		DB value	
	IRS-1A	SPOT	IRS-1A	SPOT
Training sample	9.4212	9.3343	0.5571	1.4893
1	6.8602	6.8745	0.9546	3.5146
2	7.1343	7.2301	0.9126	3.3413
3	7.3559	7.3407	0.8731	3.2078
4	8.1372	8.2166	0.779	2.8897
5	8.4162	8.4715	0.7345	2.7338

As expected, the β value is the highest and DB value is the lowest for the training set (Table 1). It is also seen that model 5 yielded superior results in terms of both the indexes, compared to other four models. The gradation of performance of five models can be established with the following β relation:

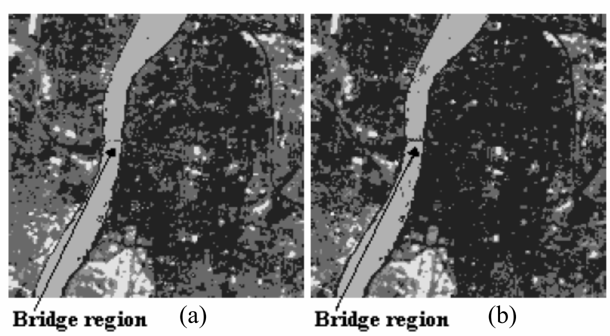
$$\beta\text{training} > \beta\text{model5} > \beta\text{model4} > \beta\text{model3} > \beta\text{model2} > \beta\text{model1} \quad (2)$$

Similar gradation of performance is also observed with DB values, which further supports the superiority of model 5.



(a) (b)

Fig. 5. Classified IRS-1A images with (a) model 1 and (b) model 5



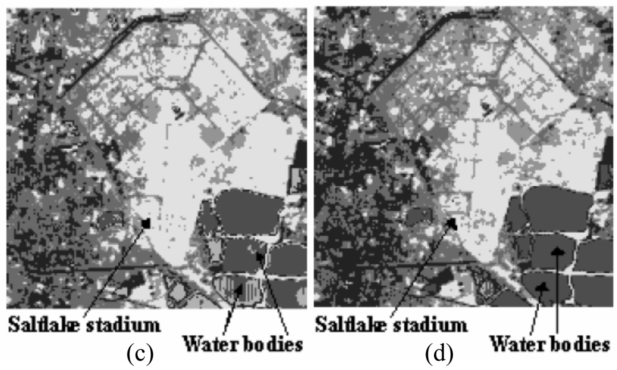
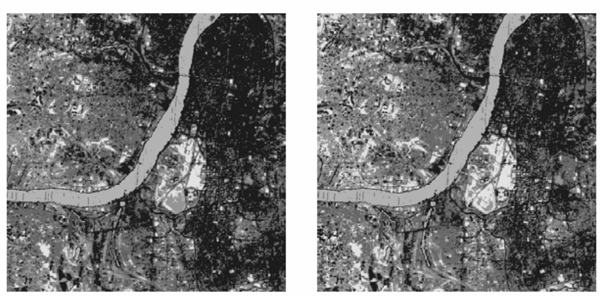


Fig. 6. (Zoomed) Two selected regions of classified IRS-1A image with (a and c) model 1, and (b and d) model 5

In order to demonstrate the significance of granular computing visually, let us consider Figs. 5a and 5b depicting the output corresponding to models 1 (without granulation) and 5 (with granulation), say for IRS-1A. It is clear from the figures that model 5 performed well in segregating different areas by properly classifying the land covers. For example, the Howrah bridge over the south part of the river is more prominent in Fig. 5b, whereas it is not so in Fig. 5a. A zoomed version of the said bridge region is shown in Figs. 6a and 6b to have an improved visualization. Similarly, the regions such as Saltlake stadium and water bodies are more distinct and well shaped with model 5 as shown in Fig. 6d (zoomed version). Similar is the case for SPOT image where Fig. 7b is superior to Fig. 7a in the sense that the different structures (e.g., roads and canals) are more prominent. These observations further justify the significance of the β and DB indexes in reflecting the performance of the models automatically without visual intervention.



(a) (b)

Fig. 7. Classified SPOT images with (a) model 1 and (b) model 5

5. Rough-Fuzzy Clustering

In the classification method, described in Section 4, f-granulation was used to model the overlapping characteristics of classes and rough sets ware considered for dimensionality reduction. In this section we describe a rough-fuzzy c-means algorithm (RFCM) which adds the concept of membership of fuzzy sets, and lower and upper approximations of rough sets into hard c-means algorithm. While the membership of fuzzy sets enables efficient handling of overlapping partitions, the rough sets deal with uncertainty, vagueness, and incompleteness in class definition [21].

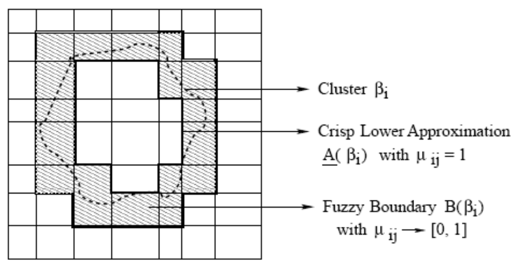


Fig. 8. Rough-fuzzy c-means: each cluster is represented by crisp lower approximations and fuzzy boundary [15, 21]

In RFCM, each cluster is represented by a centroid, a crisp lower approximation, and a fuzzy boundary (Fig. 8). The lower approximation influences the fuzziness of final partition. According to the definitions of lower approximations and boundary of rough sets, if an object belongs to lower approximations of a cluster, then the object does not belong to any other clusters. That is, the object is contained in that cluster definitely. Thus, the weights of the objects in lower approximation of a cluster should be independent of other centroids and clusters, and should not be coupled with their similarity with respect to other centroids. Also, the objects in lower approximation of a cluster should have similar influence on the corresponding centroids and cluster. Whereas, if the object belongs to the boundary of a cluster, then the object possibly belongs to that cluster and potentially belongs to another cluster. Hence, the objects in boundary regions should have different influence on the centroids and clusters. So, in RFCM, the membership values of objects in lower approximation are 1, while those in boundary region are the same as fuzzy c-means. In other word, RFCM first partitions the data into two classes - lower approximation and boundary. Only the objects in boundary are fuzzified. The new centroid is calculated based on the weighting average of the crisp lower approximation and fuzzy boundary. Computation of the centroid is modified to include the effects of both fuzzy memberships and lower and upper bounds. In essence, Rough-Fuzzy clustering tends to compromise between restrictive (hard clustering) and descriptive (fuzzy clustering) partitions.

The effectiveness of RFCM algorithm is shown, as an example, for classification of Iris data set and segmentation of brain MR images. The Iris data set is a four-dimensional data set containing 50 samples each of three types of Iris flowers. One of the three clusters (class 1) is well separated from the other two, while classes 2 and 3 have some overlap.

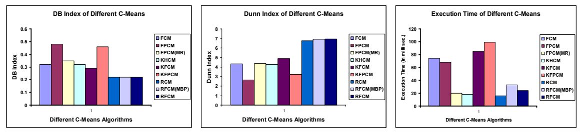


Fig. 9. Comparison of DB and Dunn Index, and execution time of HCM, FCM [16], RCM [17], $RFCM^{MBP}$ [18], and RFCM for Iris Data

The performance of different c-means algorithms is shown with respect to DB and Dunn index [20] in Fig. 9. The results reported establish the fact that RFCM provides best result having lowest DB index and highest Dunn index with lower execution time. For segmentation of brain MR images, 100 MR images with different sizes and 16 bit gray levels are tested. These MR images are collected from Advanced Medicare and Research Institute (AMRI), Kolkata, India. The comparative performance of different c-means is shown in Fig. 10 with respect to β index [19].

Some of the original images along with their segmented versions with different c-means are shown in Fig. 11. The results confirm that the RFCM algorithm produces segmented images more promising than do the conventional methods, both visually and in terms of β index.

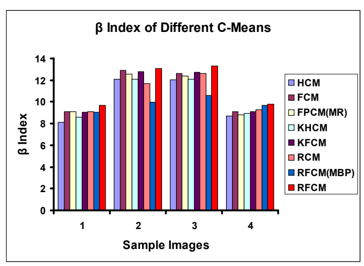
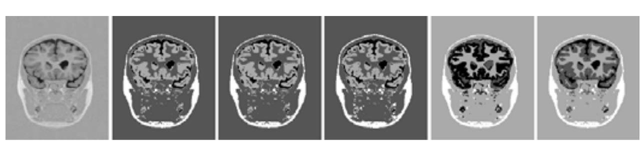


Fig. 10. Comparison of β index of HCM, FCM [16], RCM [17], RFCMMBP [18], and RFCM



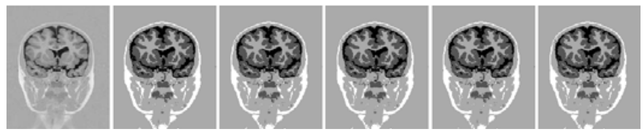


Fig. 11. Some original and segmented images of HCM, FCM [16], RCM [17], RFCM^{MBP} [18], and RFCM

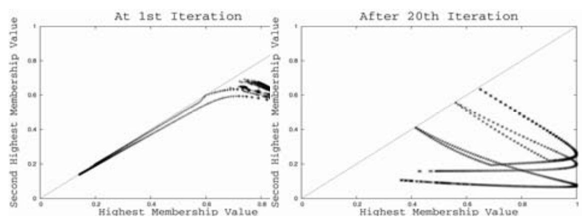


Fig. 12. Scatter plots of two highest membership values of all the objects in image data

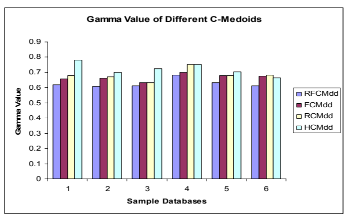
Figure 12 shows the scatter plots of the highest and second highest membership of all the objects in the data set of image at first and final iterations respectively, considering w=0.95, $(m_1 = 2.0)$ and c = 4. The diagonal line represents the zone where two highest memberships of objects are equal. It is observed that though the average differences δ between two highest memberships of the objects are very low (= 0.145) at first iteration they become ultimately very high (= 0.652) at the final iteration.

6. Rough Fuzzy C-Medoids and Amino Acid Sequence Analysis

In most pattern recognition algorithms, amino acids cannot be used directly as inputs since they are non-numerical variables. They, therefore, need encoding prior to input. In this regard, bio-basis function maps a non-numerical sequence space to a numerical feature space. It uses a kernel function to transform biological sequences to feature vectors directly. Bio-bases consist of sections of biological sequences that code for a feature of interest in the study and are responsible for the transformation of biological data to high-dimensional feature space. Transformation of input data to high-dimensional feature space is performed based on the similarity of an input sequence to a bio-basis with reference to a biological similarity matrix. Thus, the biological content in the sequences can be maximally utilized for accurate modeling. The use of similarity matrices to map features allows the bio-basis function to analyze biological sequences without the need for encoding.

One of the important issues for the bio-basis function is how to select the minimum set of bio-bases with maximum information. Here, we present an application of rough-fuzzy c-medoids (RFCMdd) algorithm [15] to select the most informative bio-bases. The objective of the RFCMdd algorithm for selection of bio-bases is to assign all amino acid subsequences to different clusters. Each of the clusters is represented by a bio-basis, which is the medoid for that cluster. The process begins by randomly choosing desired number of subsequences as the bio-bases. The subsequences are assigned to one of the clusters based on the maximum value of the similarity between the subsequence and the bio-basis. After the assignment of all the subsequences to various clusters, the new bio-bases are modified accordingly [15].

The performance of RFCMdd algorithm for bio-basis selection is presented using five whole human immunodeficiency virus (HIV) protein sequences and Cai-Chou HIV data set, which can be downloaded from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). The performance of different c-medoids algorithms such as hard c-medoids (HCMdd), fuzzy c-medoids (FCMdd) [22], rough c-medoids (RCMdd) [15], and rough-fuzzy c-medoids (RFCMdd) [15] is reported in Fig. 13 with respect to β index and γ index [15]. The results establish the superiority of RFCMdd with lowest γ index and highest β index.



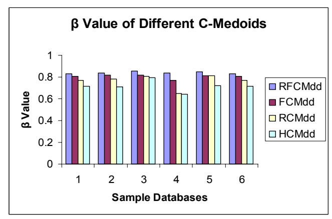


Fig. 13. Comparative performance in terms of γ and β indexes using homology alignment score

7. Generalized Rough-fuzzy Entropy and Image Ambiguity Measures

In previous examples we have demonstrated the role of granules in modeling overlapping classes, representing linguistic rules and in defining class exactness. Here we demonstrate how fuzzy boundaries of image regions, rough resemblance between nearby gray levels and rough resemblance between nearby pixels give rise to ambiguity in images, where the significance of granules in determining roughly resemblance in gray levels and pixels is evident [25]. First, we describe the concept of generalized rough sets where different cases considering both the sets and granules as crisp and/or fuzzy are explained. This is followed by entropy measures and its applications to image processing.

8. Generalized Rough Sets: Lower & Upper Approximation

The expressions for the lower and upper approximations of a set X are described for different cases, e.g., when R denotes an equivalence or a fuzzy equivalence relation and X is a crisp or a fuzzy set.

Case 1: When R denotes an equivalence relation and X is a crisp set, the expressions for the lower and upper approximations of the set X is given as

$$\underline{R}X = \{u \mid u \in U : [u]_R \subseteq X\}$$

$$\overline{R}X = \{u \mid u \in U : [u]_R \cap X \neq \varnothing\} \tag{3}$$

where $[u]_R$ denotes the granule to which the element u belongs. In this case, the pair of sets $< \underline{R}X, \overline{R}X >$ is referred to as the rough set of X and $< U, R >$ is a crisp equivalence approximation space.

Case 2: When R denotes an equivalence relation and X is a fuzzy set, the expressions for the lower and upper approximations of the set X is given as

$$\underline{R}X = \{(u, \inf_{z \in [u]_R} \mu_X(z)) \mid u \in U\}$$

$$\overline{R}X = \{(u, \sup_{z \in [u]_R} \mu_X(z)) \mid u \in U\} \tag{4}$$

where μ_X is the membership function associated with X. In this case, the pair of fuzzy sets $< \underline{R}X, \overline{R}X >$ is referred to as the rough-fuzzy set of X and $< U, R >$ is a crisp equivalence approximation space.

Case 3: Let us now consider the case when R refers to a fuzzy equivalence relation, that is, when the belongingness of every element (u) in the universe (U) to a granule $Y \in U/R$ is specified by a membership function, say m_Y, that takes values in the interval $[0,1]$ such that $\sum_Y m_Y(u) = 1$. In such a case, when X is a crisp set, the expressions for the lower and upper approximations of the set X is given as

$$\underline{R}X = \{(u, \sum_{Y \in U/R} m_Y(u) \times \inf_{\varphi \in U} \max(1 - m_Y(\varphi), C)) \mid u \in U\}$$

$$\overline{R}X = \{(u, \sum_{Y \in U/R} m_Y(u) \times \sup_{\varphi \in U} \min(m_Y(\varphi), C)) \mid u \in U\} \tag{5}$$

where

$$C = \begin{cases} 1, & \varphi \in X \\ 0, & \varphi \notin X \end{cases}$$

In the above, the symbols $\sum$ (sum) and $\times$ (product) respectively represent specific fuzzy union and intersection operations. Note that, one may consider any fuzzy union and intersection operation instead of the sum and product operations by judging their suitability with respect to the underlying application. The pair of fuzzy sets $< \underline{R}X, \overline{R}X >$ is referred to as the fuzzy rough set of X in this case and $< U, R >$ is a fuzzy equivalence approximation space.

Case 4: In Case 3 of R referring to a fuzzy equivalence relation, when X is a fuzzy set, the expressions for the lower and upper approximations of the set X is given as

$$\underline{R}X = \{(u, \sum_{Y \in U/R} m_Y(u) \times \inf_{\varphi \in U} \max(1 - m_Y(\varphi), \mu_X(\varphi))) \mid u \in U\}$$

$$\overline{R}X = \{(u, \sum_{Y \in U/R} m_Y(u) \times \sup_{\varphi \in U} \min(m_Y(\varphi), \mu_X(\varphi))) \mid u \in U\} \quad (6)$$

The pair of fuzzy sets $< \underline{R}X, \overline{R}X >$ is referred as the fuzzy rough-fuzzy set of X and $< U, R >$ is again a fuzzy equivalence approximation space. From the above explanation, it is obvious that the set of expressions in cases 1 -3 are special cases of the set of expressions for the lower and upper approximations given in Case 4. Pictorial diagram of lower and upper approximations for Case 4 is shown in Fig 14.

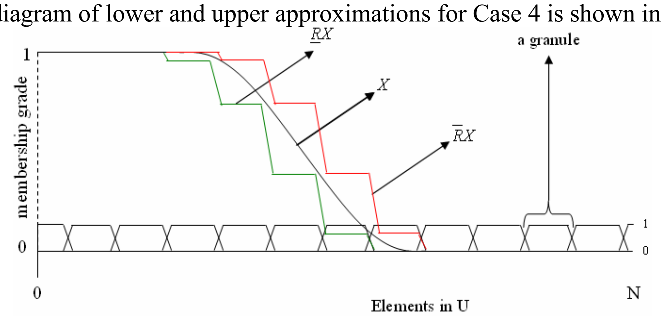


Fig. 14. The pair $< \underline{R}X, \overline{R}X >$ is referred to as the fuzzy rough-fuzzy set of X

Entropy Measures.

Let us now provide two classes of entropy measures based on roughness measures of a set and its complement in order to quantify the incompleteness of knowledge about a universe. One of them is based on logarithmic gain function, defined as

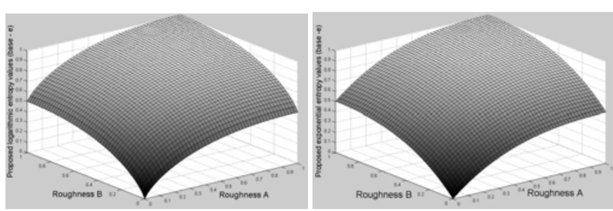
$$H_R^L(X) = -\frac{1}{2}[\rho_R(X)\log_\beta(\frac{\rho_R(X)}{\beta}) + \rho_R(X^c)\log_\beta(\frac{\rho_R(X^c)}{\beta})] \quad (7)$$

where β denotes the base of the logarithmic function used and $X \subseteq U$ stands for the complement of the set X in the universe. The various entropy measures of this class are obtained by calculating the roughness values $\rho_R(X) = 1 - \frac{|\underline{R}X|}{|\overline{R}X|}$ and $\rho_R(X^c) = 1 - \frac{|\underline{R}X^c|}{|\overline{R}X^c|}$ considering the different ways of obtaining the lower and upper approximations of the vaguely definable set X. Note that, the 'gain in incompleteness' term is taken as $-\log_\beta(\frac{\rho_R}{\beta})$ in (7) and for $\beta > 1$ it takes a value in the interval $[1, \infty]$. The other class of entropy measures, as obtained by considering an exponential function to measure the 'gain in incompleteness', is:

$$H_R^E(X) = \frac{1}{2}[\rho_R(X)\beta^{(1-\rho_R(X))} + \rho_R(X^c)\beta^{(1-\rho_R(X^c))}] \quad (8)$$

where β denotes the base of the exponential function used. Similar to the class of entropy measures H_R^L, the various entropy measures of this class are obtained by using the different ways of obtaining the lower and upper approximations of X in order to calculate $\rho_R(X)$ and $\rho_R(X^c)$. The 'gain in incompleteness' term is taken as $\beta^{(1-\rho_R)}$ in (2) and for $\beta > 1$ it takes a value in the finite interval $[1, \beta]$.

In Figure 15, the values of H_R^L and H_R^E are shown for all possible values of the roughness measures $\rho_R(X)$ and $\rho_R(X^c)$ considering $\beta = e$. Figure 16 shows the plots of the proposed entropies for different values of β, when $\rho_R(X) = \rho_R(X^c)$.



(a) Logarithmic (b) Exponential

Fig. 15. Plots of the proposed classes of entropy measures for various roughness values $A = \rho_R(X)$ and $B = \rho_R(X^c)$

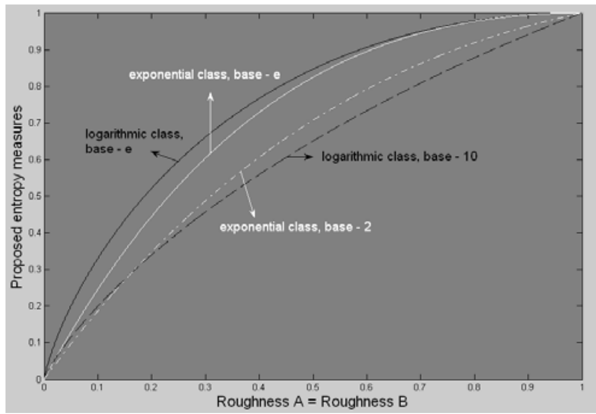


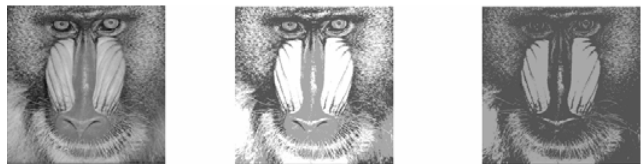
Fig. 16. Proposed entropy measures for a few different β values, when

$$\rho_R(X) = \rho_R(X^c)\ (A = B)$$

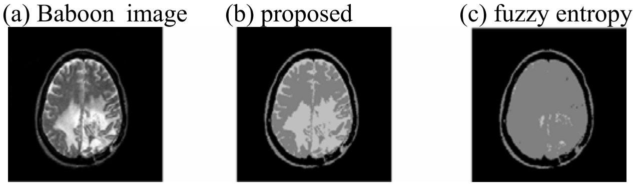
9. Image Ambiguity measures and Segmentation:

Using the aforesaid entropy definitions, we compute grayness and spatial ambiguity measures of an image. Grayness ambiguity refers to indefiniteness associated with deciding whether a pixel or a clump of pixels (granule) is white or black. That is, it concerns with the indefiniteness due to fuzziness as well as granularity in gray values. Spatial Ambiguity, on the other hand, refers to indefiniteness in shape and geometry of various regions where indefiniteness is concerned with both intensity and spatial location of individual pixel or group of pixels. These ambiguity measures can be minimized by changing the cross-over point of the membership function to find a set of minima corresponding to different thresholds of an image.

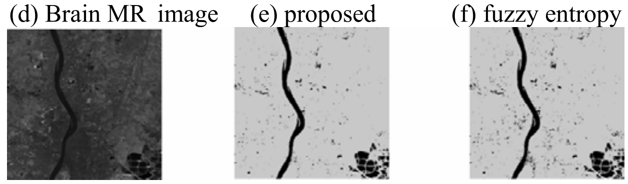
Fig 17 shows the segmentation results of three images, as an example, using grayness ambiguity measures based on rough-fuzzy entropy (proposed) and fuzzy entropy [26]. In the former case, membership of a pixel is dependent on the granule (defined over one-dimensional gray scale) to which it belongs, and it is independent of its spatial location. In the latter case, the membership of a pixel is entirely dependent on its own gray value, and it is independent of its spatial location. Therefore the improvement in segmentation results by rough-fuzzy entropy as compared to fuzzy entropy in Fig 17 is due to the effect of granules. The same is quantitatively demonstrated in Fig. 18 for 45 other images where β-index for segmentation is seen in almost all cases to be higher for outputs corresponding to rough-fuzzy entropy.



(a) Baboon image (b) proposed (c) fuzzy entropy



(d) Brain MR image (e) proposed (f) fuzzy entropy



(g) Remote sensing image (h) proposed (i) fuzzy entropy

Fig. 17. Comparative segmentation results

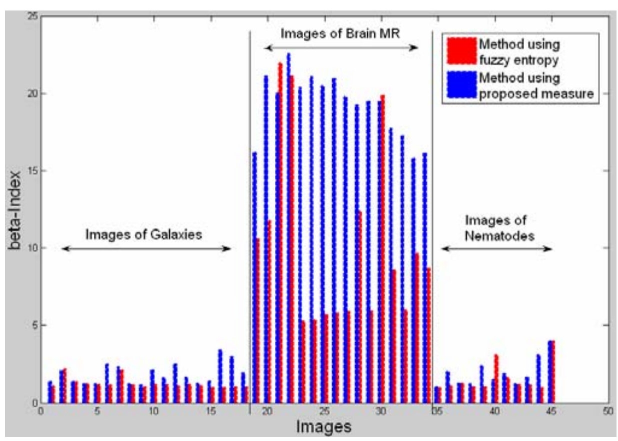


Fig. 18. β-index for segmentation results on 45 images (Significance of adding the concept of granules to gray level fuzziness is evident)

10. Conclusion

Rough-fuzzy approach in soft computing frame work is considered for pattern recognition and knowledge discovery. The concept of knowledge encoding using rough sets and the role *f*-granulation to make it more efficient are illustrated. Examples of judicious integration for tasks like rough-fuzzy case generation, rough-fuzzy classification, rough-fuzzy c-medoids and rough-fuzzy entropy measures with their merits and characteristics are described. Cases generated are efficient for dimensionality reduction, classification and retrieval. Merits of rough-fuzzy clustering in protein sequence analysis for determining bio-bases and segmentation of brain MR images are demonstrated. C-medoids can be used for any other relational clustering. Other applications of rough-fuzzy computing in bioinformatics and medical imaging are available in a recent book [30]. Class dependent granulation with neighborhood rough sets is seen to have better class discrimination ability than class independent granulation with Pawlak's rough sets. The algorithm is very useful even when the number of available training samples is small. The said concept has recently been extended to form the theory of rough-wavelet granulation [31]. The effect of granularity in improving the quality of image segmentation *vis-à-vis* fuzzy entropic segmentation is adequately established. The same can be demonstrated for other image processing operations. The entropy expressions can be used to define other kinds of measures for efficient image processing [32]. The concept of the aforesaid rough-fuzzy computing has a significant role in modeling the computational theory of perception (CTP). Readers may refer in this context to some pioneering articles of Zadeh and others [33-38]

It may further be mentioned here that granulation is inherent in human thinking and reasoning processes. Granulation is also a computing paradigm, among others such as self-reproduction, self-organization, functioning of brain, perception, Darwinian evolution, group behavior, cell membranes, and morphogenesis that are abstracted from natural phenomena. A good survey on natural computing explaining its different facets is provided in [39]. Therefore the present article also enriches the literature in natural computational framework.

Acknowledgement. The author acknowledges J.C. Bose Fellowship of the Govt. of India, and his colleagues / students Drs. Pabitra Mitra, Pradipta Maji, Saroj Meher and Debashis Sen for their contributions in different phases.

References

[1] A. Pal and S. K. Pal, Pattern recognition: Evolution of methodologies and data mining, In *Pattern Recognition: From Classical to Modern Approaches*, Eds., S. K. Pal and A. Pal, World Scientific, Singapore, 2002, pp. 1-23.

[2] S. Mitra, S. K. Pal and P. Mitra, Data mining in soft computing framework: A survey, *IEEE Trans. Neural Networks*, vol. 13, no. 1, pp. 3-14, 2002.

[3] L. A. Zadeh, A new direction in AI: Toward a computational theory of perceptions, *AI Magazine*, vol. 22, pp. 73-84, 2001.
[4] L. A. Zadeh, Foreword, *Neuro-Fuzzy Pattern Recognition: Methods in Soft Computing*, (Authors: S. K. Pal and S. Mitra), Wiley, New York, 1999.
[5] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer, Dordrecht, 1991.
[6] S. K. Pal and A. Skowron, Eds., *Rough-Fuzzy Hybridization: A New Trend in Decision Making*, Springer-Verlag, Singapore, 1999.
[7] S.K. Pal and J. F. Peters, Eds., *Rough-fuzzy Image Analysis: Foundations and Methodologies*, Chapman & Hall/CRC, Boca Raton, Florida, 2010.
[8] S. K. Pal and A. Skowron, Eds., Special issue on Rough Sets, Pattern Recognition and Data Mining, *Pattern Recognition Letters*, vol. 24, no. 6, 2003.
[9] J. L. Kolodner, *Case-Based Reasoning*, Morgan Kaufmann, San Mateo, CA, 1993.
[10] S. K. Pal, T.S. Dillon, and D. S. Yeung, Eds., *Soft Computing in Case Based Reasoning*, Springer, London, 2001.
[11] S. K. Pal and S. C. K. Shiu, *Foundations of Soft Case Based Reasoning*, John Wiley, NY, 2003.
[12] S. K. Pal and P. Mitra, Case generation using rough sets with fuzzy discretization, *IEEE Trans. Knowledge and Data Engineering*, vol. 16, no. 3, pp. 292-300, 2004.
[13] S. K. Pal and P. Mitra, *Pattern Recognition Algorithms for Data Mining*, Chapman & Hall CRC Press, Boca Raton, FL, 2004.
[14] Y. Li, S. C. K. Shiu and S. K. Pal, Combining feature reduction and case selection in building CBR classifiers, *IEEE Trans. on Knowledge and Data Engineering*, vol. 18, no. 3, pp. 415-429, 2006.
[15] P. Maji and S. K. Pal, Rough-fuzzy C-medoids algorithm and selection of bio-basis for amino acid sequence analysis, *IEEE Trans. Knowledge and Data Engineering,* vol. 19, no. 6, pp. 859-872, 2007.
[16] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, NY, 1981.
[17] P. Lingras and C. West, Interval set clustering of web users with rough k-means, *Journal of Intelligent Information Systems*, vol. 23, no. 1, pp. 5-16, 2004.
[18] S. Mitra, H. Banka, and W. Pedrycz, Rough-fuzzy collaborative clustering, *IEEE Trans. on Systems, Man, and Cybernetics - Part B: Cybernetics*, vol. 36, pp. 795-805, 2006.
[19] S. K. Pal, A. Ghosh, and B. U. Sankar, Segmentation of remotely sensed images with fuzzy thresholding, and quantitative evaluation, *Int. Journal of Remote Sensing*, vol. 2, no. 11, pp. 2269-2300, 2000.
[20] J. C. Bezdek and N. R. Pal, *Some* new indexes for cluster validity, *IEEE Trans. on System, Man, and Cybernetics, Part B*, vol. 28, pp. 301-315, 1988.
[21] P. Maji and S. K. Pal, Rough set based generalized fuzzy C-means algorithm and quantitative indices, *IEEE Trans. on System, Man and Cybernetics, Part B*, vol. 37, no. 6, pp. 1529--1540, 2007.
[22] R. Krishnapuram, A. Joshi, O. Nasraoui, and L. Yi, Low complexity fuzzy relational clustering algorithms for web mining, *IEEE Trans. on Fuzzy System*, vol. 9, pp. 595-607, 2001.
[23] S. Bandyopadhyay and S. K. Pal, *Classification and Learning Using Genetic Algorithms: Applications in Bioinformatics and Web Intelligence*, Springer, Berlin, 2007.
[24] P. Maji and S. K. Pal, Feature selection using f-information measures in fuzzy approximation spaces, *IEEE Trans. Knowledge and Data Engineering*, vol. 22, no. 6, pp. 854-867, 2010.
[25] D. Sen and S. K. Pal, Generalized rough sets, entropy and image ambiguity measures, *IEEE Transactions on System, Man and Cybernetics - Part B*, vol.39, no.1, pp.117-128, 2009.

[26] S. K. Pal, R.A. King and A.A. Hashim, Automatic gray level thresholding through index of fuzziness and entropy, *Pattern Recognition Letters*, vol. 1, pp.141-146, 1983.
[27] S. K. Pal, S. K. Meher and S. Dutta, Pattern classification using class-dependent rough-fuzzy granular space, *Proc. Fifth Int. Conf. on Rough Set and Knowledge Technology*, Beijing, China, LNAI, vol. 6401, pp. 54-61, 2010.
[28] Q. Hu, D. Yu, J. Liu and C. Wu, Neighborhood rough set based heterogeneous feature subset selection, *Inform. Sciences,* vol. 178, pp. 3577-3594, 2008.
[29] S. K. Pal, Computational theory of perception (CTP), rough-fuzzy uncertainty analysis and mining in bioinformatics and web intelligence: A unified framework, *LNCS Trans. on Rough Sets,* LNCS vol. 5946, pp. 106-129, 2009.
[30] P. Maji and S.K. Pal, *Rough-fuzzy Pattern Recognition: Applications in Bioinformatics and Medical Imaging,* John Wiley, N.Y., (to appear)
[31] S. K. Meher and S. K. Pal, Rough-wavelet granular space and classification of multispectral remote sensing image. *Applied Soft Computing* (to appear)
[32] D. Sen and S. K. Pal, Histogram thresholding using fuzzy and rough measures of association errors, *IEEE Transactions on Image Processing*, vol.18, no.4, pp.879-888, 2009.
[33] L.A. Zadeh, Fuzzy sets and information granularity. In: M. Gupta, R. Ragade and R. Yager (Eds.), *Advances in Fuzzy Set Theory and Applications*, North-Holland Publishing Co., Amsterdam, pp.3-18, 1979.
[34] L.A. Zadeh, Fuzzy logic, neural networks, and soft computing. *Communications of the ACM*, vol.37, pp.77-84, 1994.
[35] L.A. Zadeh, Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets and Systems*, vol.19, pp.111-127, 1997.
[36] L.A. Zadeh, A new direction in AI: Towards a computational theory of perceptions. *AI Magazine,* vol. 22, no.1, pp.73-84, 2001.
[37] L.A. Zadeh, Generalized theory of uncertainty (GTU) - principal concepts and ideas, *Computational Statistics \& Data Analysis*, vol.51, no.1, pp.15—46, 2006.
[38] A. Skowron and P. Wasilewski, An Introduction to Perception Based Computing, In: T.-H. Kim, Y.-H. Lee, B.H. Kang, D. Slcezak (Eds.) *Proceedings of the Second International Conference on Future Generation Information Technology, FGIT 2010*, Jeju Island, Korea, December 13-15, 2010, *Lecture Notes in Computer Science*, no.6485, pp.12-25, 2010.
[39] L. Kari and R. Rozenberg, The many facets of natural computing, *Communications of the ACM*, vol. 51, no.10, pp.72-83, 2008.

Sankar K. Pal
Center for Soft Computing Research,
Indian Statistical Institute,
Kolkata 700108, India,
sankar@isical.ac.in

The Notion of Consistency in the Presence of Knowledge Operator

Saswati Sadhu

This work deals with the notion of consistency of sets of formulae involving knowledge operator. Adoption of the following procedure may be beneficial to investigate the possibility of accommodating knowledge of a sentence and knowledge of its negation together in the same system. We confine ourselves to Jaakko Hintikka's system of knowledge as outlined in his seminal work 'Knowledge and Belief' published in 1962 [1]. For the general notions of propositional modal logic, we have referred to [2, 3]

Notions of consistency and consequence in the classical propositional logic

The notion of consistency is defined in terms of derivability in classical propositional calculus, which leads to the obvious question whether it is possible to have a definition the other way round and whether these two notions are equivalent, in some sense, to each other. Stanislaw J. Surma has given a satisfactory answer to this question in his paper [4] treating the notion of consequence as an operator (Cn) on subsets of formulae and consistency being dealt with as a unary predicate ($\mathscr{H}$-$Cons$) on the subsets of formulae in the following manner: S is the set of all well formed formulae of classical propositional logic, $\sim$ and $\rightarrow$ being the primitive logical connectives. ($\mathscr{S}$-$Cons$) satisfies the following conditions:

($\mathscr{S}$-**Cons 1**) $\mathscr{S}\text{-}Cons \subseteq \mathscr{P}(S)$
($\mathscr{S}$-**Cons 2**) if $X \subseteq Y \in \mathscr{S}\text{-}Cons$, then $X \in \mathscr{S}\text{-}Cons$
($\mathscr{S}$-**Cons 3**) if $X \subseteq S$ and $X \notin \mathscr{S}\text{-}Cons$, then there is a finite subset Y in X such that $Y \notin \mathscr{S}\text{-}Cons$,
($\mathscr{S}$-**Cons 4**) if $x \in S$ and $X \in \mathscr{S}\text{-}Cons$, then either $X \cup \{x\} \in \mathscr{S}\text{-}Cons$ or $X \cup \{\sim x\} \in \mathscr{S}\text{-}Cons$,
($\mathscr{S}$-**Cons 5**) if $x \in S$, then $\{x, \sim x\} \notin \mathscr{S}\text{-}Cons$,
($\mathscr{S}$-**Cons 6**) if $x, y \in S$, then $X \cup \{x, \sim y\} \in \mathscr{S}\text{-}Cons$ if and only if $X \cup \{\sim (x \rightarrow y)\} \in \mathscr{S}\text{-}Cons$.

With the help of the above conditions it can be shown that the set $\{x \in S : \{\sim x\} \notin \mathscr{S}\text{-}Cons\}$ coincides with the set of all theorems of the two-valued zero order logic.

Considering $S, \sim, \rightarrow$ as before, Cn, a function from the power set $\mathscr{P}(S)$ of S into itself, satisfies the following conditions:

(**Cn 1**) if $X \subseteq S$, then $Cn(X) \subseteq S$,
(**Cn 2**) if $X \subseteq S$, then $X \subseteq Cn(X)$,

(**Cn 3**) if $X \subseteq S$, then $Cn(Cn(X)) \subseteq Cn(X)$,
(**Cn 4**) if $X \subseteq Y \subseteq S$, then $Cn(X) \subseteq Cn(Y)$,
(**Cn 5**) if $X \subseteq S$, then $Cn(X) \subseteq \cup\{Cn(Y) : Y$ is finite and $Y \subseteq X\}$,
(**Cn 6**) if $X \subseteq S$ and $x, y \in S$, then $x \to y \in Cn(X)$ if and only if $y \in Cn(X \cup \{x\})$,
(**Cn 7**) if $x \in S$, then $Cn(\{x, \sim x\}) = S$,
(**Cn 8**) if $x \in S$, then $Cn(\{x\}) \cap Cn(\{\sim x\}) = Cn(\emptyset)$.

According to (**Cn 1**) – (**Cn 4**), Cn can be treated as a closure operator in S. It can be proved that the condition (**Cn 6**) is satisfied by the operator Cn if and only if the set $Cn(\emptyset)$ coincides with the set of all theorems of the implication logic in the sense of Hilbert. It can also be proved that the conditions (**Cn 6**) – (**Cn 8**) are satisfied by the operator Cn if and only if $Cn(\emptyset)$ coincides with the set of all theorems of the two-valued logic with negation and implication.

After axiomatizing these two notions, Surma, in his paper, has defined Cn in terms of $\mathscr{S}$-*Cons* and vice-versa as follows:
$Cn(X) = \{x \in S : X \cup \{x\} \in \mathscr{S}\text{-}Cons\}$ — (1)
$\mathscr{S}\text{-}Cons = \{X \subseteq S : Cn(X) \neq S\}$ — (2)

In either case $\mathscr{S}$-*Cons* satisfies the conditions ($\mathscr{S}$-*Cons* 1) to ($\mathscr{S}$-*Cons* 6) and Cn satisfies the conditions (Cn 1) to (Cn 8). An $\mathscr{S}$-*Cons* defined, using (2), from a Cn which is again defined, using (1), from another $\mathscr{S}$-*Cons* are the same and the case is similar for Cn. In other words the given two sets of axioms are equivalent to each other.

Consistency and consequence extended

Jaakko Hintikka in his work 'Knowledge and Belief' presented the notion of consistency of propositions involving knowledge operator, through some rules, stated below:

(**A**.**PKK***) If a set X of sentences is consistent and if $Kx_1 \in X,\ Kx_2 \in X, ..., Kx_n \in X,\ Py \in X$, then the set $\{Kx_1, Kx_2, ..., Kx_n, y\}$ is also consistent.
(**A**.**PK***) If a set X of sentences is consistent and if $Kx_1 \in X,\ Kx_2 \in X, ..., Kx_n \in X,\ Py \in X$, then the set $\{x_1, x_2, ..., x_n, y\}$ is also consistent.
(**A**.**K**) If a set X of sentences is consistent and if $Kx \in X$, then the set $X \cup \{x\}$ is also consistent.
(**A**. $\sim$ **K**) If a set X of sentences is consistent and if $\sim Kx \in X$, then the set $X \cup \{P \sim x\}$ is also consistent.
(**A**. $\sim$ **P**) If a set X of sentences is consistent and if $\sim Px \in X$, then the set $X \cup \{K \sim x\}$ is also consistent.

In the above rules Kx is interpreted as 'a particular agent knows that x' and Px as 'it is possible for all that a particular agent knows, that x'.

Consistency rules introduced by Hintikka, which do not refer to the operators K or P are as follows:

($\mathbf{A}.\wedge$) If a set X of sentences is consistent and if $x \wedge y \in X$, then the set $X \cup \{x, y\}$ is also consistent.
($\mathbf{A}.\vee$) If a set X of sentences is consistent and if $x \vee y \in X$, then the set $X \cup \{x\}$ or the set $X \cup \{y\}$ is consistent(or both are).
($\mathbf{A}.\sim$) If X is a set of sentences and if $x \in X$ and $\sim x \in X$, then the set X is inconsistent.
($\mathbf{A}.\sim\wedge$) If a set X of sentences is consistent and if $\sim (x \wedge y) \in X$, then the set obtained from X by replacing $\sim (x \wedge y)$ by $\sim x \vee \sim y$ is also consistent.
($\mathbf{A}.\sim\vee$) If a set X of sentences is consistent and if $\sim (x \vee y) \in X$, then the set obtained from X by replacing $\sim (x \vee y)$ by $\sim x \wedge \sim y$ is also consistent.
($\mathbf{A}.\sim\sim$) If a set X of sentences is consistent and if $\sim\sim x \in X$, then the set obtained from X by replacing $\sim\sim x$ by x is also consistent.

From the rules proposed by Hintikka it is not clear if it would be possible to derive some obviously necessary results. To achieve this, the set of axioms proposed by Hintikka has been modified in this work in the following way:
The connectives $\vee, \rightarrow, \leftrightarrow$ are defined in the usual manner and Px is defined as $\sim K \sim x$.
Hintikka's notion of consistency can be treated as a unary predicate ($\mathscr{H}$-$Cons$) on the subsets of formulae of propositional modal logic satisfying the following modified set of axioms:

($\mathscr{H}$-**Cons 1**) $\mathscr{H}\text{-}Cons \subseteq \mathscr{P}(S)$,
($\mathscr{H}$-**Cons 2**) if $X \subseteq Y \in \mathscr{H}\text{-}Cons$, then $X \in \mathscr{H}\text{-}Cons$,
($\mathscr{H}$-**Cons 3**) if $X \subseteq S$ and $X \notin \mathscr{H}\text{-}Cons$, then there is a finite subset Y in X such that $Y \notin \mathscr{H}\text{-}Cons$,
($\mathscr{H}$-**Cons 4**) if $x \in S$ and $X \in \mathscr{H}\text{-}Cons$, then either $X \cup \{x\} \in \mathscr{H}\text{-}Cons$ or $X \cup \{\sim x\} \in \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\wedge$) if $X \in \mathscr{H}\text{-}Cons$ and $x \wedge y \in X$, then $X \cup \{x, y\} \in \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\wedge\,\mathbf{conv}$) if $X \in \mathscr{H}\text{-}Cons$ and $x, y \in X$, then $X \cup \{x \wedge y\} \in \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\sim$) if $X \subseteq S$ and $x, \sim x \in X$, then $X \notin \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\sim\sim$) if $X \in \mathscr{H}\text{-}Cons$ and $\sim\sim x \in X$, then $(X \setminus \{\sim\sim x\}) \cup \{x\} \in \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\sim\sim\,\mathbf{conv}$) if $X \in \mathscr{H}\text{-}Cons$ and $x \in X$, then $(X \setminus \{x\}) \cup \{\sim\sim x\} \in \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\mathbf{PKK}^*$) if $X \in \mathscr{H}\text{-}Cons$ and $Kx_1,\ Kx_2, ..., Kx_n,\ Py \in X$, then $\{Kx_1, Kx_2, ..., Kx_n, y\} \in \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\mathbf{PK}^*$) if $X \in \mathscr{H}\text{-}Cons$ and $Kx_1,\ Kx_2, ..., Kx_n,\ Py \in X$, then $\{x_1, x_2, ..., x_n, y\} \in \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\mathbf{K}$) if $X \in \mathscr{H}\text{-}Cons$ and $Kx \in X$, then $X \cup \{x\} \in \mathscr{H}\text{-}Cons$,
($\mathbf{A}.\sim\mathbf{K}$) if $X \in \mathscr{H}\text{-}Cons$ and $\sim Kx \in X$, then $X \cup \{P \sim x\} \in \mathscr{H}\text{-}Cons$.

In this work it has been shown that Surma's notion of consistency ($\mathscr{S}$-$Cons$) and consequence (Cn) may be extended upto S, the set of all well formed formulae of propositional modal logic where $\sim, \rightarrow$ are the primitive logical connectives and K is the primitive modal operator, by adding the following axioms:

($\mathscr{S}$-**Cons K**) if $x, y \in S$, then $\{K(x \rightarrow y), Kx, \sim Ky\} \notin \mathscr{S}\text{-}Cons$,
($\mathscr{S}$-**Cons T**) if $x \in S$, then $\{Kx, \sim x\} \notin \mathscr{S}\text{-}Cons$,
($\mathscr{S}$-**Cons S4**) if $x \in S$, then $\{Kx, \sim KKx\} \notin \mathscr{S}\text{-}Cons$,
($\mathscr{S}$-**Cons Nec**) if $x \in S$ and $\{\sim x\} \notin \mathscr{S}\text{-}Cons$, then $\{\sim Kx\} \notin \mathscr{S}\text{-}Cons$.

and

(**Cn K**) if $x, y \in S$, then $Ky \in Cn(\{K(x \rightarrow y), Kx\})$,
(**Cn T**) if $x \in S$, then $x \in Cn(\{Kx\})$,
(**Cn S4**) if $x \in S$, then $KKx \in Cn(\{Kx\})$,
(**Cn Nec**) if $x \in Cn(\emptyset)$, then $Kx \in Cn(\emptyset)$.

Making use of definitions (1) and (2) in the extended notion it is obvious that the extended sets of axioms for $\mathscr{S}$-$Cons$ and Cn are equivalent to each other. It has also been established that Hintikka's notion of consistency (modified) is equivalent to Surma's notion of consistency, extended upto modal system S_4, and consequently the equivalence between Surma's notion of consequence extended upto modal system S_4 and Hintikka's notion of consistency (modified) is obtained, the equivalence being in the same sense as that in the classical case.

Notes

This work was done under the guidance of Professor Mihir Kumar Chakraborty. There are some gaps in this work, such as

– From the rules proposed by Hintikka it has not been clear if it would be possible to derive some obviously necessary results e.g.
  - if a set of sentences are consistent and if $x \in X$ and $y \in X$, then the set $X \cup \{x \wedge y\}$ is also consistent,
  - if a set X of sentences is consistent and if $x \in X$, then the set obtained from X by replacing x by $\sim\sim x$ is also consistent,
  - if a set X of sentences is consistent and if $\sim x \vee \sim y \in X$, then the set obtained from X by replacing $\sim x \vee \sim y$ by $\sim (x \wedge y)$ is also consistent,
  - if a set X of sentences is inconsistent, then there is a finite subset Y in X such that Y is inconsistent, and so on.

– In order to prepare the modified set of axioms for $\mathscr{H}$-$Cons$ the following axioms have been deleted from Hintikka's system, which are derivable from the modified set:

  (**A**.$\vee$) if $X \in \mathscr{H}\text{-}Cons$ and $x \vee y \in X$, then $X \cup \{x\} \in \mathscr{H}\text{-}Cons$ or $X \cup \{y\} \in \mathscr{H}\text{-}Cons$ or both.

($\mathbf{A}.\sim\wedge$) if $X \in \mathscr{H}\text{-}Cons$ and $\sim(x \wedge y) \in X$, then $(X \setminus \{\sim(x \wedge y)\}) \cup \{\sim x \vee \sim y\} \in \mathscr{H}\text{-}Cons$.
($\mathbf{A}.\sim\vee$) if $X \in \mathscr{H}\text{-}Cons$ and $\sim(x \vee y) \in X$, then $(X \setminus \{\sim(x \vee y)\}) \cup \{\sim x \wedge \sim y\} \in \mathscr{H}\text{-}Cons$.
($\mathbf{A}.\sim\mathbf{P}$) $X \in \mathscr{H}\text{-}Cons$ and $\sim Px \in X$, then $X \cup \{K \sim x\} \in \mathscr{H}\text{-}Cons$.

Whether the original set of Hintikka axioms is equivalent to the modified set has not been investigated.

References

1. J. Hintikka: *Knowledge and Belief*, Cornell University Press, Ithaka, 1962.
2. G.E. Hughes and M.J. Cresswell: *An Introduction to Modal Logic*, Methuen & Co. Ltd., London, 1968.
3. G.E. Hughes and M.J. Cresswell: *A Companion to Modal Logic*, Methuen, London, 1984.
4. S.J. Surma: The Growth of Logic out of the Foundational Research in Mathematics, *Modern Logic - A Survey* (ed. E. Agazzi), D. Reidel Publishing Company, pp. 15-33, 1980.

Saswati Sadhu
Christopher Road Govt. Spon. H.S. School (for Girls),
Kolkata, India.
saswatisadhu@gmail.com

Axiomatization of Topological Quasi Boolean Algebra

Anirban Saha

Topological quasi-Boolean algebra (tqBa) is an algebraic structure which was first considered during the study of rough sets. TqBa was primarily discussed in [1, 11]. The formal definition of tqBa can be described as follows.

An abstract algebra $\mathcal{A} =< A, \wedge, \vee, \sim, I, 0, 1 >$ is said to be a *topological quasi-Boolean algebra* (*tqBa*) if and only if

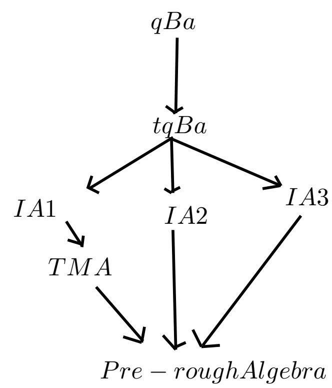
1. $\mathcal{A} =< A, \wedge, \vee, \sim, 0, 1 >$ is a quasi-Boolean algebra (qBa) [8] i.e., $\mathcal{A} =< A, \wedge, \vee, \sim, 0, 1 >$ is a distributive lattice with least element 0 and greatest element 1 satisfying
   (a) $\sim\sim a = a$ for all $a \in A$
   (b) $\sim (a \vee b) =\sim a \wedge \sim b$ for all $a, b \in A$
2. The unary operator I satisfies the following properties for all $a, b \in A$
   (a) $I(a \wedge b) = Ia \wedge Ib$,
   (b) $Ia \wedge a = a$,
   (c) $IIa = Ia$,
   (d) $I1 = 1$,
   (e) $CIa = Ia$, where $Ca \equiv\sim I \sim Ia$.

With the axioms of tqBa if the following three additional axioms

[2(e)] $\sim Ia \vee Ia = 1$,
[2(f)] $I(a \vee b) = Ia \vee Ib$,
[2(g)] $Ia \wedge Ib = Ia$ and $Ca \wedge Cb = Ca$ imply $a \wedge b = a$

are taken then another new algebraic structure is obtained, which is called a *pre-rough algebra* [2]. Now with the axioms of tqBa if the axiom 2(e) is added then also a new algebraic structure can be obtained – this is called the *intermediate algebra 1* (IA1). Similarly *intermediate algebra 2* (IA2) and *intermediate algebra 3* (IA3) can be obtained by adding the axioms 2(f) and 2(g) respectively to the axioms of tqBa [10]. Again, with the axioms of intermediate algebra I if the axiom $\sim Ia \wedge a =\sim a \wedge a$ is taken, then another important algebraic structure can be obtained. This is called the *tetravalent modal algebra* (TMA). This algebra was first considered by A. Monteiro, motivated by L. Monteiro's independence proof of an axiomatization of three-valued Łukasiewicz algebras [6]. Later, I. Loureiro, the last student of A. Monteiro, carried out a brief study on TMA [4, 5].
It can be shown that every pre-rough algebra is a TMA but not conversely.

Relations of these algebras are shown below.



$X \to Y$: X with some additional axiom(s) gives Y

As far as the study of the logics corresponding to these algebras is concerned, it can be found that sequent calculi of the tqBa, IA1, IA2, IA3 and pre-rough algebra are already defined [10]. The Hilbert system of the logic corresponding to the pre-rough algebra is also obtained [2], where the implication '$\Rightarrow$' of the logic is defined as
$\alpha \Rightarrow \beta \equiv (\sim I\alpha \vee I\beta) \wedge (\sim C\alpha \vee C\beta)$.

In case of TMA, there is a four-valued Monteiro propositional calculus, the algebraic counterpart of which is TMA [3, 4]. Two implications '$\to$, $\Rightarrow$' of the logic are defined as
$\alpha \to \beta \equiv C \sim \alpha \vee \beta$ and
$\alpha \Rightarrow \beta \equiv (\alpha \to \beta) \wedge (C\alpha \to C(\alpha \wedge \beta))$.

On the other hand, another important aspect in the context of obtaining a Hilbert system for a class of algebras is the finite axiomatization of that class. One can find that the class of TMAs is finitely axiomatizable [5]. From the study of the tqBas so far, it can be shown that some of the subclasses of the class of tqBas are finitely axiomatizable [7, 9]. Our attempt now is to investigate the tqBa in these directions, in order to establish a Hilbert-style axiomatization for the corresponding logic.

References

1. M. Banerjee and M.K. Chakraborty: Rough algebra, *Bull Polish Acad Sc (Math)*, 41(4):293-297, 1993.
2. M. Banerjee and M.K. Chakraborty: Rough sets through algebraic logic, *Fundamenta Informaticae* 28(3-4): 211-221, 1996.
3. E. Bianco: Fourvalued Monteiro propositional calculus *XII Latin American Symposium on Mathematical Logic. Abstracts of Contributed Papers*, San Jose, Costa Rica 3, 2004.
4. I. Loureiro: Finitely generated free tetravalent modal algebras, *Discrete Mathematics* 46 (1983), pp. 41-48.
5. I. Loureiro: Principal Congruences of Tetravalent Modal Algebras, *Notre Dame Journal of Formal Logic* 26(1),76-80, 1985.

6. L. Monteiro: Axiomes independants pour les algebres de Lukasiewicz trivalentes, *Bulletin de la Society des Sciences Mathimatiques et Physiques de la R. P. Roumanie*, Nouvelle Sirie 7, 199-202, 1963.
7. K. Pal: Modal Logic and Topological Quasi-Boolean Algebras *Masters' Thesis*, Dept. of Mathematics and Statistics, Indian Institute of Technology, Kanpur, 2005.
8. H. Rasiowa: *An Algebraic Approach to Non-Classical Logics*, North-Holland Publishing Company, 1994.
9. A. Saha: A Step Towards a Hilbert-Style Axiomatic System for Topological Quasi-Boolean Algebras, *Phd Course work (Logic) project*, Dept. of Pure Mathematics, University of Calcutta, April 2011.
10. J. Sen and M.K. Chakraborty: A study of interconnections between rough and Łukasiewicz 3-valued logic, *Fundamenta Informaticae* 51: 311 324, 2002.
11. A. Wasilewska and M. Banerjee: Rough sets and topological quasi- Boolean algebras, *Proc. 23rd Ann. ACM CSC'95 Rough Sets and Database Mining*, in: T.Y. Lin (ed.), San Jose State Univ., 121-128, 1995.

Anirban Saha
Supreme Knowledge Foundation Group of Institutions,
Hooghly, West Bengal.
anirbansaha31182@gmail.com

Generalized Rough Sets, Implication Lattices and Consistency-Degree

Pulak Samanta

Abstract. Indiscernibility relation enhanced by partition on an universe is known to be knowledge. If the Indiscernibility relation fails to be transitive then the granules of knowledge overlap and they form a covering of the universe. An extensive survey of rough set theory based on partition, covering and relation is done. Implication lattice / Implication diagram with respect to approximation pairs is shown. Dependency degree of knowledges based on partition as well as covering have been considered. Consistency degree between two knowledges and inconsistency measure in both the cases are obtained.

1 Introduction

I am compiling here the work I have done so far with Professor Mihir Kumar Chakraborty, the references are [11, 26–28].

Pawlak introduced Rough Set Theory in the year 1982 by considering an approximation space $< U, R >$, where U is a non-empty set and R is an equivalence relation on U. So R forms a partition on U. Given any subset A of U, the lower and upper approximations $\underline{A}_R$ and $\overline{A}^R$ are then defined by $\underline{A}_R = \{x|\ [x]_R \subseteq A\}$ and $\overline{A}^R = \{x|\ [x]_R \cap A \neq \phi\}$ where $[x]_R$ is the equivalence class of x with respect to R. We can immediately observe that the following properties of lower and upper approximations hold.

(1a) $\underline{U} = U$ (Co-normality) (1b) $\overline{U} = U$ (Co-normality)
(2a) $\underline{\phi} = \phi$ (Normality) (2b) $\overline{\phi} = \phi$ (Normality)
(3a) $\underline{A} \subseteq A$ (Contraction) (3b) $A \subseteq \overline{A}$ (Extension)
(4a) $\underline{A \cap B} = \underline{A} \cap \underline{B}$ (Multiplication) (4b) $\overline{A \cup B} = \overline{A} \cup \overline{B}$ (Addition)
(5a) $\underline{(\underline{A})} = \underline{A}$ (Idempotency) (5b) $\overline{(\overline{A})} = \overline{A}$ (Idempotency)
(6) $\underline{(\sim A)} = \sim (\overline{A})$, $\overline{(\sim A)} = \sim (\underline{A})$ (Duality)
(7a) $A \subseteq B \Rightarrow \underline{A} \subseteq \underline{B}$ (Monotone) (7b) $A \subseteq B \Rightarrow \overline{A} \subseteq \overline{B}$ (Monotone)
(8a) $A \subseteq \underline{(\overline{A})}$ (8b) $\overline{(\underline{A})} \subseteq A$

For addressing more generalized approaches we shall have to expand this list of properties by splitting the equalities 4a, 4b, 5a, 5b into two inequalities. There are at least three distinct approaches towards generalization viz. Relational approach, Covering based approach and Operator based approach.

Dependency degree between two knowledges given by two partitions on a set defined by Novotný and Pawlak [15, 17–19]. Based on this we defined Dependency degree based on covering [26], Consistency degree between two knowledges based on partition [11, 26], and also based on covering [26].

2 Various types of lower and upper approximations

2.1 Covering-based operators

The references of our relevant papers for this section is [26–28]. Let A be a subset of U. Here is a list of constructions of various kinds of the lower-upper approximations of A with respect to a covering.

We have used $\underline{P}_i, \overline{P}^i$ $i = 1, 2, 3, 4$ to acknowledge Pomykala, since to our knowledge he first studied the lower and upper approximations with respect to covering. However, the pair $< \underline{P}_4, \overline{P}^4 >$ was due to Pawlak but looking apparently different. $\underline{C}_i, \overline{C}^i,\ i = 1, 2, 3, 4, 5$ are other covering based approximations which are essentially duals. $\underline{C}$ and $\overline{C}$ with extra symbols are also covering based lower and upper approximation operators, the symbols being taken from the respective papers straightway. This group of pairs barring $\underline{C}_{Gr}$ and $\overline{C}^{Gr}$ are non-duals.

$\underline{P}_1(A) = \{x : N_x^{\mathcal{C}} \subseteq A\}$
$\overline{P}^1(A) = \cup\{C_i : C_i \cap A \neq \phi\}$ [23, 26, 31]

$\underline{P}_2(A) = \cup\{N_x^{\mathcal{C}} : N_x^{\mathcal{C}} \subseteq A\}$
$\overline{P}^2(A) = \{z : \forall y(z \in N_y^{\mathcal{C}} \Rightarrow N_y^{\mathcal{C}} \cap A \neq \phi)\}$ [23, 26]

$\underline{P}_3(A) = \cup\{C_i : C_i \subseteq A\}$
$\overline{P}^3(A) = \{y : \forall C_i(y \in Ci \Rightarrow C_i \cap A \neq \phi)\}$ [13, 23, 26, 29, 31, 35]

$\underline{P}_4(A) = \cup\{P_x^{\mathcal{C}} : P_x^{\mathcal{C}} \subseteq A\}$
$\overline{P}^4(A) = \cup\{P_x^{\mathcal{C}} : P_x^{\mathcal{C}} \cap A \neq \phi\}$ [5, 11, 13, 19, 23, 25, 26, 29, 30, 33, 35]

$\underline{C}_1(A) = \cup\{C_i : C_i \in \mathcal{C}, C_i \subseteq A\}$
$\overline{C}^1(A) = \sim \underline{C}_1(\sim A) = \cap\{\sim C_i : C_i \in \mathcal{C}, C_i \cap A = \phi\}$ [25]

$\underline{C}_2(A) = \{x \in U : N(x) \subseteq A\}$
$\overline{C}^2(A) = \{x \in U : N(x) \cap A \neq \phi\}$ [13, 25]

$\underline{C}_3(A) = \{x \in U : \exists u(u \in N(x) \wedge N(u) \subseteq A)\}$
$\overline{C}^3(A) = \{x \in U : \forall u(u \in N(x) \rightarrow N(u) \cap A \neq \phi)\}$ [25]

$\underline{C}_4(A) = \{x \in U : \forall u(x \in N(u) \rightarrow N(u) \subseteq A)\}$
$\overline{C}^4(A) = \cup\{N(x) : N(x) \cap A \neq \phi\}$ [25]

$\underline{C}_5(A) = \{x \in U : \forall u(x \in N(u) \rightarrow u \in A)\}$
$\overline{C}^5(A) = \cup\{N(x) : x \in A\}$ [25]

With the same lower approximation there are a few different upper approximations. In the following we have borrowed the symbols from corresponding authors.

$\underline{C}_*(A) = \underline{C}_-(A) = \underline{C}_\#(A) = \underline{C}_@(A) = \underline{C}_+(A) = \underline{C}_\%(A)$
$= \cup\{C_i \in \mathcal{C} : C_i \subseteq A\} \equiv \underline{P}_3(A)$ [14]

$\overline{C}^*(A) = \underline{C}_*(A) \cup \{Md(x) : x \in A \setminus A_*\}$ [14, 35]

$\overline{C}^-(A) = \cup\{C_i : C_i \cap A \neq \phi\}$ [14]

$\overline{C}^\#(A) = \cup\{Md(x) : x \in A\}$ [14, 35]

$\overline{C}^@(A) = \underline{C}_@(A) \cup \{C_i : C_i \cap (A \setminus \underline{C}_@(A)) \neq \phi\}$ [14]

$\overline{C}^+(A) = \underline{C}_+(A) \cup \{Neighbour(x) : x \in A \setminus \underline{C}_+(A)\}$ [14, 33]

$\overline{C}^\%(A) = \underline{C}_\%(A) \cup \{\sim \cup\{Friends(y) : x \in A \setminus \underline{C}_\%(A), y \in e.f(x)\}\}$ [14]

Another type of lower and upper approximations is defined with the help of covering.

Let, $Gr_*(A) = \cup\{C_i \in \mathcal{C} : C_i \subseteq A\} \equiv \underline{P}_3(A)$.

This is taken as lower approximation of A and is denoted by $\underline{C}_{Gr}(A)$.

$Gr^*(A) = \cup\{C_i \in \mathcal{C} : C_i \cap A \neq \phi\} \equiv \overline{P}^1(A)$.
The upper approximation is defined by $\overline{C}^{Gr}(A) = Gr^*(A) \setminus NEG_{Gr}(A)$, where, $NEG_{Gr}(A) = \underline{C}_{Gr}(\sim A)$, $\sim A$ being the complement of A [29].

The following table shows the properties satisfied by the respective lower-upper approximation pairs. It may be noted that the properties in the list given in the introduction is now expanded because of obvious reasons.

Table I

Properties of covering based approximations

	P_1	P_2	P_3	P_4	C_1	C_2	C_3	C_4	C_5	C_{Gr}	C_*	C_-	$C_\#$	$C_@$	C_+	$C_\%$
Dual	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N
$\underline{\phi} = \phi = \overline{\phi}$	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y$_*$	Y	Y
$\underline{U} = U = \overline{U}$	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y$_*$	Y	Y
$\underline{A \cap B} \subseteq \underline{A} \cap \underline{B}$	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y$_*$	Y	Y
$\underline{A} \cap \underline{B} \subseteq \underline{A \cap B}$	Y	N$_*$	N$_*$	Y	N$_*$	Y	N$_*$	Y	Y	N	N$_*$	N$_*$	N$_*$	N$_*$	N$_*$	N
$\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$	Y	N$_*$	N$_*$	Y	N$_*$	Y	N$_*$	Y	Y	N	N$_*$	Y	N$_*$	Y$_*$	Y	N
$\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y$_*$	Y	N
$A \subseteq B \Rightarrow \underline{A} \subseteq \underline{B}$	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y$_*$	Y	Y
$A \subseteq B \Rightarrow \overline{A} \subseteq \overline{B}$	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N$_*$	Y	Y
$\underline{A} \subseteq$ A	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y$_*$	Y	Y
$A \subseteq \overline{A}$	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y$_*$	Y	N
$\underline{A} \subseteq \overline{A}$	Y	Y	Y	Y	Y	Y	N$_*$	Y	Y	Y	Y	Y	Y	Y$_*$	Y	Y
$A \subseteq \underline{(\overline{A})}$	Y$_*$	N$_*$	Y$_*$	Y	N$_*$	N	N$_*$	Y	N	N$_*$	Y$_*$	Y$_*$	Y$_*$	Y$_*$	N$_*$	N$_*$
$\overline{(\underline{A})} \subseteq A$	Y$_*$	N$_*$	Y$_*$	Y	N$_*$	N$_*$	N$_*$	Y	N$_*$	N$_*$	N$_*$	N$_*$	Y$_*$	Y$_*$	Y$_*$	N$_*$
$\underline{A} \subseteq \underline{(\underline{A})}$	N	Y	Y	Y	Y	Y	N$_*$	N	Y	Y	Y	Y	Y	Y$_*$	Y	Y
$\overline{(\overline{A})} \subseteq \overline{A}$	N	Y	Y	Y	Y	Y	N$_*$	N	Y	Y	Y	N	N	Y$_*$	Y	Y
$\overline{A} \subseteq \underline{(\overline{A})}$	N$_*$	N$_*$	N$_*$	Y	N$_*$	N$_*$	N$_*$	N	N$_*$	N$_*$	Y$_*$	Y	Y$_*$	Y$_*$	N$_*$	N$_*$
$\overline{(\underline{A})} \subseteq \underline{A}$	N$_*$	N$_*$	N$_*$	Y	N$_*$	N$_*$	N$_*$	N$_*$	N$_*$	N$_*$	N$_*$	N$_*$	Y$_*$	Y$_*$	Y$_*$	N$_*$

In table **I**, Y means 'yes, the property holds' and N means 'no, the property does not hold'.

Properties that we have verified ourselves are marked $*$. Other results are taken straightway from the respective papers.

Remark 1. From C_* to $C_\%$ the lower approximations as stated earlier are the same as that of P_3 viz. $\underline{P}_3(A) = \cup\{C_i : C_i \subseteq A\}$. $\overline{P}^3(A)$ is its dual. Naturally all other upper approximations viz. $\overline{C}^*(A)$ to $\overline{C}^\%(A)$ are non duals. Now an attempt may be made to retain these upper approximations and take their respective duals as corresponding lower approximations. One such pair is suggested by Cattaneo [7]. He takes $\overline{C}^{-}(A)$ as the upper approximation and its dual the lower one. His motivation is completely mathematical making a connection with pre-topological (Čech) behavior.

Remark 2. In the same paper [7], Cattaneo has taken another interesting mathematical (Tarski Topological) approach to define lower and upper approximations in terms of pseudo-open and pseudo-closed sets. Let $\mathcal{C} = \{C_i\}$ be a covering of the universe U. A set $A \subseteq X$ is called pseudo-open if A is the union of some sets in $\mathcal{C}$ and is called pseudo-closed if A is the intersection of the complements of some sets in $\mathcal{C}$. Then $\underline{A} = \{P \subseteq A : P$ is pseudo-open$\}$ and $\overline{A} = \cap\{Q \subseteq A : Q$ is pseudo-closed$\}$. We can check that the lower and upper approximations are $\underline{P}_3(A)$ and $\overline{P}^3(A)$ respectively. Thus, this pair of operators gets an elegant topological interpretation.

Remark 3. We have come across some other covering based operators [32] which virtually coincide with some of those presented here.

2.2 Relation-based operators

The reference of our relevant paper for this section is [28]. We are now presenting the table corresponding to above mentioned properties but further expanded, with respect to relation based definition of lower-upper approximations. The 2nd and 3rd rows are now split into four rows, reasons for which will be clear from the table. Here we have used R for any relation, r, s, t denote reflexivity, symmetry and transitivity. R with suffix(es) means that the relation possesses the corresponding property or properties. There are other important conditions that may be ascribed to R e.g. seriality or Archimedeanness, but in this study we are not considering them.

No originality is claimed in table **II**. (Almost) all the results are nicely furnished in Yao's and Zhu's papers [30, 31, 34]. Besides, anybody familiar with elementary modal logic will recognize R as the accessibility relation and lower and upper approximation operators as the semantic counterparts of necessity and possibility operators.

Table II

Properties of relation based approximations

	R	R_r	R_s	R_t	R_{rs}	R_{rt}	R_{st}	R_{rst}
Dual	Y	Y	Y	Y	Y	Y	Y	Y
$\underline{\phi} = \phi$	N	N	N	N	Y	Y	N	Y
$\phi = \overline{\phi}$	Y	Y	Y	Y	Y	Y	Y	Y
$\underline{U} = U$	Y	Y	Y	Y	Y	Y	Y	Y
$U = \overline{U}$	N	N	N	N	Y	Y	N	Y
$\underline{A \cap B} \subseteq \underline{A} \cap \underline{B}$	Y	Y	Y	Y	Y	Y	Y	Y
$\underline{A} \cap \underline{B} \subseteq \underline{A \cap B}$	Y	Y	Y	Y	Y	Y	Y	Y
$\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$	Y	Y	Y	Y	Y	Y	Y	Y
$\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$	Y	Y	Y	Y	Y	Y	Y	Y
$A \subseteq B \Rightarrow \underline{A} \subseteq \underline{B}$	Y	Y	Y	Y	Y	Y	Y	Y
$A \subseteq B \Rightarrow \overline{A} \subseteq \overline{B}$	Y	Y	Y	Y	Y	Y	Y	Y
$\underline{A} \subseteq A$	N	Y	N	N	Y	Y	N	Y
$A \subseteq \overline{A}$	N	Y	N	N	Y	Y	N	Y
$\underline{A} \subseteq \overline{A}$	N	Y	N	N	Y	Y	N	Y
$A \subseteq \underline{(\overline{A})}$	N	N	Y	N	Y	N	Y	Y
$\overline{(\underline{A})} \subseteq A$	N	N	Y	N	Y	N	Y	Y
$\underline{A} \subseteq \underline{(\underline{A})}$	N	N	N	Y	N	Y	Y	Y
$\overline{(\overline{A})} \subseteq \overline{A}$	N	N	N	Y	N	Y	Y	Y
$\overline{A} \subseteq \underline{(\overline{A})}$	N	N	N	N	N	N	N	Y
$\overline{(\underline{A})} \subseteq \underline{A}$	N	N	N	N	N	N	N	Y

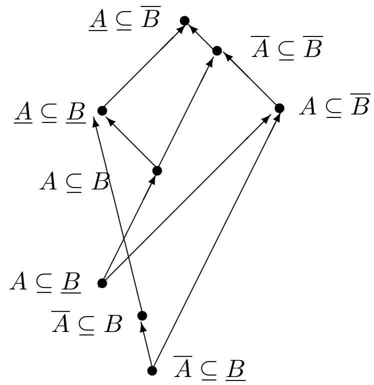
2.3 Operator based Approach

The reference of our relevant paper for this section is [28]. This is also called axiomatic approach [30] in which two operators $\_$ and $^-$ are defined on $\mathcal{P}(U)$, the powerset of U satisfying certain axioms which are some properties like those of the list given at the beginning or its expanded version. That means lower and upper approximations $\underline{A}$ and $\overline{A}$ of a set A are taken straightway satisfying certain axioms. This approach is important mathematically and logically since in whatever way $\underline{A}$ and $\overline{A}$ might have been defined what matters here are the properties that they possess. The entire group of algebras / topologies / logics that have emerged from rough set approach care only for these properties. These in turn has pushed towards another kind of generalization of rough sets viz. the abstract algebraic and topological approaches [1–3, 5, 8, 16, 20–22, 24].

3 Partial ordering of inclusion relations and implication lattices

The references of our relevant papers for this section is [27, 28]. Given two subsets A, B of the universe U there are nine possible inclusions $P \subseteq Q$, $P \in \{\underline{A}, A, \overline{A}\}$, $Q \in \{\underline{B}, B, \overline{B}\}$. In case the lower and upper approximations arise out of a partition on U, that is the systems P_4 or R_{rst}, we have the following equivalences, $\{\underline{A} \subseteq \overline{B}\}$, $\{\underline{A} \subseteq \underline{B}, \underline{A} \subseteq B\}$, $\{A \subseteq \overline{B}, \overline{A} \subseteq \overline{B}\}$, $\{A \subseteq B\}$ and $\{A \subseteq \underline{B}, \overline{A} \subseteq \underline{B}, \overline{A} \subseteq B\}$ in the sense that inclusions belonging to the same group are equivalent that is, each implies the other. This implication relation is then extended for the equivalence classes of relations. It may be observed that with respect to this latter implication, the set of above equivalence classes in most cases form a lattice which has been called an implication lattice. The equivalence classes of inclusion relations are disjoint but one class may imply the other. This implication is defined by an arrow ($\rightarrow$) in the diagrams. Implication lattices were first introduced in [10]. In this more general context we shall see that implication relation between clusters of inclusions does not form a lattice in general - it forms a partial order relation. Among the 14 different diagrams, here are three examples of the above mentioned diagrams.
For a categorization of inclusion relation in these systems, properties that are responsible are the following :
(i) $A \subseteq B \Rightarrow \underline{A} \subseteq \underline{B}$ (ii) $A \subseteq B \Rightarrow \overline{A} \subseteq \overline{B}$ (iii) $\underline{A} \subseteq A$
(iv) $A \subseteq \overline{A}$ (v) $\underline{A} \subseteq \overline{A}$ (vi) $A \subseteq \underline{(\overline{A})}$
(vii) $\overline{(\underline{A})} \subseteq A$ (viii) $\underline{A} \subseteq \underline{(\underline{A})}$ (ix) $\overline{(\overline{A})} \subseteq \overline{A}$
(x) $\overline{A} \subseteq \underline{(\overline{A})}$
The arrow indicates that the node at the tail implies that at the head. The properties used and equivalence classes formed with more than one element are shown by the side of the diagrams. The other equivalence classes are singletons. In the diagrams a representative of each class is depicted without using brackets.

Diagram 1 (System $C_{\%}$)

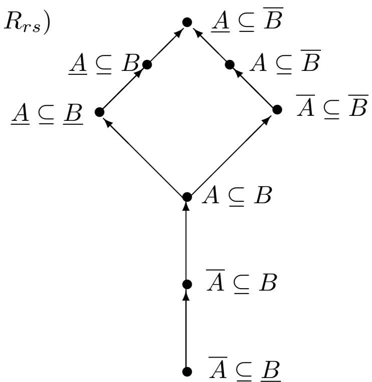


Properties used :

(i), (ii), (iii), (v), (viii) and (ix).

Equivalence classes :
$\{\underline{A} \subseteq \underline{B}, \underline{A} \subseteq B\}$
and singletons

Diagram 2 (Systems P_1 and C_4 and R_{rs})

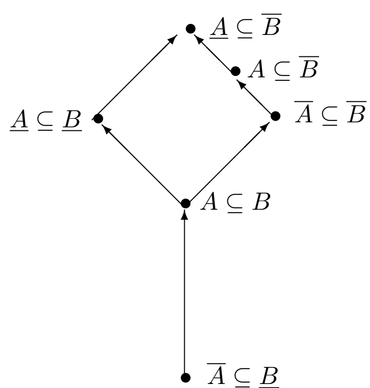


Properties used :

(i) to (iv), (vi) and (vii).

Equivalence classes :
$\{\overline{A} \subseteq B, A \subseteq \underline{B}\}$
and singletons

Diagram 3 (Systems P_4 and R_{rst})



Properties used :

(i) to (iv), (vi) to (viii) and (x).

Equivalence classes :
$\{\underline{A} \subseteq \underline{B}, \underline{A} \subseteq B\}$, $\{A \subseteq \overline{B}, \overline{A} \subseteq \overline{B}\}$
and $\{A \subseteq \underline{B}, \overline{A} \subseteq \underline{B}, \overline{A} \subseteq B\}$
and singletons

4 Dependency of knowledge

4.1 Dependency of knowledge based on partition

The reference of our relevant paper for this section is [11]. We would accept the basic philosophy that a knowledge of an agent about an universe is her ability to categorize objects inhabiting it through information received from various sources or perception in the form of attribute-value data. For this section we start with the indiscernibility relation caused by the attribute-value system. So, knowledge is defined as follows.

Definition 1. *Knowledge : A knowledge is a pair,* $< U, P >$ *where* U *is a non-empty finite set and* P *is an equivalence relation on* U. P *will also denote the partition generated by the equivalence relation.*

Definition 2. *Finer and Coarser Knowledge : A knowledge* P *is said to be finer than the knowledge* Q *if every block of the partition* P *is included in some block of the partition* Q. *In such a case* Q *is said to coarser than* P. *We shall write it as* $P \preceq Q$.

We recall a few notions due to Pawlak (and others) e.g. P-positive region of Q and based upon it dependency-degree of knowledges.

Definition 3. *Let* P *and* Q *be two equivalence relations over* U. *The* P*-positive region of* Q*, denoted by* $Pos_P(Q)$ *is defined by*
$Pos_P(Q) = \bigcup_{X \in U/Q} \underline{P}X$ *, where* $\underline{P}X = \bigcup\{Y \in U/P : Y \subseteq X\}$ *called* P*-lower approximation of* X.

Definition 4. *Dependency degree : Knowledge* Q *depends in a degree* k $(0 \leq k \leq 1)$ *on knowledge* P *, written as* $P \Rightarrow_k Q$*, iff* $k = \frac{CardPos_P(Q)}{CardU}$ *where Card denotes cardinality of the set.*
If $k = 1$ *, we say that* Q *totally depends on* P *and we write* $P \Rightarrow Q$*; and if* $k = 0$ *we say that* Q *is totally independent of* P.

Viewing from the angle of multi-valuedness one can say that the sentence 'The knowledge Q depends on the knowledge P' instead of being only 'true'(1) or 'false'(0) may receive other intermediate truth-values, the value k being determined as above. This approach justifies the term 'partial dependency' as well.

In propositions 1,2 and 3, we enlist some elementary, often trivial, properties of dependency degree, most of which are present in [15, 19]. Some of these properties e.g. proposition 3(v) will constitute the basis of definitions and results of the next section.

Proposition 1. *(i)* $[x]_{P_1 \cap P_2} = [x]_{P_1} \cap [x]_{P_2}$,
(ii) If $P \Rightarrow Q$ *and* $R \preceq P$ *then* $R \Rightarrow Q$,
(iii)If $P \Rightarrow Q$ *and* $Q \preceq R$ *then* $P \Rightarrow R$,
(iv)If $P \Rightarrow Q$ *and* $Q \Rightarrow R$ *then* $P \Rightarrow R$,
(v)If $P \Rightarrow R$ *and* $Q \Rightarrow R$ *then* $P \cap Q \Rightarrow R$,

(vi) If $P \Rightarrow R \cap Q$ then $P \Rightarrow R$ and $P \Rightarrow Q$,
(vii) If $P \Rightarrow Q$ and $Q \cap R \Rightarrow T$ then $P \cap R \Rightarrow T$,
(viii) If $P \Rightarrow Q$ and $R \Rightarrow T$ then $P \cap R \Rightarrow Q \cap T$.

Proposition 2. *(i) If $P' \preceq P$ then $\underline{P'}X \supseteq \underline{P}X$,*
(ii) If $P \Rightarrow_a Q$ and $P' \preceq P$ then $P' \Rightarrow_b Q$ where $b \geq a$,
(iii) If $P \Rightarrow_a Q$ and $P \preceq P'$ then $P' \Rightarrow_b Q$ where $b \leq a$,
(iv) If $P \Rightarrow_a Q$ and $Q' \preceq Q$ then $P \Rightarrow_b Q'$ where $b \leq a$,
(v) If $P \Rightarrow_a Q$ and $Q \preceq Q'$ then $P \Rightarrow_b Q'$ where $a \leq b$.

Proposition 3. *(i) If $R \Rightarrow_a P$ and $Q \Rightarrow_b P$ then $R \cap Q \Rightarrow_c P$ for some $c \geq Max(a, b)$,*
(ii) If $R \cap P \Rightarrow_a Q$ then $R \Rightarrow_b Q$ and $P \Rightarrow_c Q$ for some $b, c \leq a$,
(iii) If $R \Rightarrow_a Q$ and $R \Rightarrow_b P$ then $R \Rightarrow_c Q \cap P$ for some $c \leq Min(a, b)$,
(iv) If $R \Rightarrow_a Q \cap P$ then $R \Rightarrow_b Q$ and $R \Rightarrow_c P$ for some $b, c \geq a$,
(v) If $R \Rightarrow_a P$ and $P \Rightarrow_b Q$ then $R \Rightarrow_c Q$ for some $c \geq a + b - 1$.

4.2 Dependency of knowledge based on covering

The reference of our relevant paper for this section is [26]. A covering $\mathcal{C}$ of a set U is a collection of subsets $\{C_i\}$ of U such that $\cup C_i = U$. It is often important to define a knowledge in terms of covering and not by partition which is a special case of covering. Given a covering $\mathcal{C}$ one can define a binary relation $R^{\mathcal{C}}$ on U, which is a tolerance relation (reflexive, symmetric): $xR^{\mathcal{C}}y$ holds iff $x, y \in C_i$ for some i, where the set $\{C_i\}$ constitute the covering.

Definition 5. *A tolerance space is a structure $S = < U, R >$, where U is a nonempty set of objects and R is a reflexive and symmetric binary relation defined on U.*

A tolerance class of a tolerance space $< U, R >$ is a maximal subset of U such that any two elements of it are mutually related.

In the context of knowledge when the indiscernibility relation R is only reflexive and symmetric (and not necessarily transitive) the approximation system $< U, R >$ is a tolerance space. In such a case the granules of the Knowledge may be formed in many different ways.

Definition 6. *[4, 9] A covering is said to be genuine covering if $C_i \subseteq C_j$ implies $C_i = C_j$.*

For any genuine covering $\mathcal{C}$ it is immediate that the elements of $\mathcal{C}$ are all tolerance classes of the relation $R^{\mathcal{C}}$.

Definition 7. *We shall say that a covering $\mathcal{C}_1$ is finer than a covering $\mathcal{C}_2$ written as $\mathcal{C}_1 \preceq \mathcal{C}_2$ iff $\forall C'_j \in \mathcal{C}_2\ \exists\ C_{j1}, C_{j2}, ..., C_{jn}$ such that $C'_j = C_{j1} \cup C_{j2} \cup ... \cup C_{jn}$ where, $C_{j1}, C_{j2}, ..., C_{jn} \in \mathcal{C}_1$ i.e. every element of $\mathcal{C}_2$ may be expressed as the union of some elements of $\mathcal{C}_1$.*

Let R be a tolerance relation in U. Then the family $\mathcal{C}(R)$ of all tolerances classes of R is a covering of U. The pair $(U, \mathcal{C})$ will be called generalized approximation space, where U is a set and $\mathcal{C}$ is a covering of U. We shall however assume U to be finite in the sequel.

Proposition 4. *If $\mathcal{C}_1 \preceq \mathcal{C}_2$ then $P_1 \preceq P_2$ where P_1, P_2 are the partitions corresponding to $\mathcal{C}_1$, and $\mathcal{C}_2$ respectively.*

Definition 8. *We define $\mathcal{C}_1$-Positive region of $\mathcal{C}_2$ as $Pos_{\mathcal{C}_1}\mathcal{C}_2 = \bigcup_{X \in \mathcal{C}_2} \underline{\mathcal{C}_1}(X)$.*

Definition 9. *Dependency degree with respect to covering : $\mathcal{C}_1$ depends in a degree k $(0 \leq k \leq 1)$ on $\mathcal{C}_2$, written as $\mathcal{C}_1 \Rightarrow_k \mathcal{C}_2$, iff $k = \frac{|Pos_{\mathcal{C}_1}\mathcal{C}_2|}{|U|}$ where $|X|$ denotes cardinality of the set X. We shall also write $k = Dep(\mathcal{C}_1, \mathcal{C}_2)$.*
If $k = 1$, $\mathcal{C}_1$ is said to be totally dependent on $\mathcal{C}_2$ and we write $\mathcal{C}_1 \Rightarrow \mathcal{C}_2$; and if $k = 0$ we say that $\mathcal{C}_2$ is totally independent of $\mathcal{C}_1$.

5 Consistency of knowledge

We begin this section by mentioning that [11, 26] are the relevant references of our papers. Two knowledges P and Q on U where P and Q are partitions may be considered as fully consistent if and only if $U/P = U/Q$, that is P,Q generate exactly the same granules. This is equivalent to $P \Rightarrow Q$ and $Q \Rightarrow P$. So, a natural measure of consistency degree of P and Q might be the truth-value of the non-classical sentence "Q depends on $P \wedge P$ depends on Q" computed by a suitable conjunction operator applied on the truth-values of the two component sentences Thus a binary predicate $Cons$ may be created such that $Cons(P, Q)$ will stand for the above conjunctive sentence. A triangular norm (or t-norm) used in fuzzy-literature and many-valued logic scenario is a potential candidate for computing $\wedge$. A t-norm is a mapping $t : [0, 1] \to [0, 1]$ satisfying (i) $t(a, 1) = a$, (ii) $b \leq d$ implies $t(a, b) \leq t(a, d)$, (iii) $t(a, b) = t(b, a)$, (iv) $t(a, t(b, d)) = t(t(a, b), d)$. It follows that $t(a, 0) = 0$. Typical examples of t-norm are : $min(a, b)$ (Gödel), $max(0, a + b - 1)$ (Lukasicwicz), $a \times b$ (Godo,Hajek).
These are conjunction operators used extensively and are in some sense the basic t-norms [12]. With $1 - x$ as negation operator the De-Morgan dual of t-norms called s-norms are obtained as $s(a, b) = 1 - t(1 - a, 1 - b)$. Values of disjunctive sentences are computed by s-norms.

There is however a difficulty in using a t-norm in the present context. We would like to have the following assumptions to hold.

Assumption 1: Knowledges P,Q shall be fully consistent iff they generate the same partition.

Assumption 2: Knowledges P,Q shall be fully inconsistent iff no granule generated by one is contained in any granule generated by the other.

The translation of the above demands in mathematical terms is that the conjunction operator $\star$ should fulfill the conditions :

$\star(a, b) = 1$ iff $a = 1, b = 1$
and $\star(a, b) = 0$ iff $a = 0, b = 0$.

No t-norm satisfies the second. So we define consistency degree as follows:

Definition 10. *Let P and Q be two knowledges such that $P \Rightarrow_a Q$ and $Q \Rightarrow_b P$. The consistency degree between the two knowledges denoted by $Cons(P, Q)$ is given by $Cons(P, Q) = \frac{a+b+nab}{n+2}$, where n is a non negative integer.*

Definition 11. *Two knowledges P and Q are said to be fully consistent if $Cons(P, Q) = 1$.*
Two knowledge P and Q are said to be fully inconsistent if $Cons(P, Q) = 0$.

Although any choice of n satisfies the initial requirements, some special values for it may be of special significance e.g $n = 0$, $n = Card(U)$. We shall refer'n' as the 'consistency constant' or simply 'constant' in the sequel.

If the t-norm is taken to be $max(0, a+b-1)$, then the corresponding s-norm is $min(1, a + b)$. For the t-norm $min(a, b)$, the s-norm is $max(a, b)$. There is an order relation in the t-norms / s-norms, viz.
any t-norm $\leq min \leq max \leq$ any s-norm. In particular,
$max(0, a + b - 1) \leq min(a, b) \leq max(a, b) \leq min(1, a + b)$.

Where does the Cons function situate itself in this chain - might be an interesting and useful query. The following proposition answers this question.

Proposition 5. *$max(0, a + b - 1) \leq Cons(P, Q) \leq max(a, b)$ if $P \Rightarrow_a Q$ and $Q \Rightarrow_b P$.*

To compare $Cons(P, Q)$ and $min(a, b)$, we have,

Proposition 6. *Let P and Q be two knowledges and $P \Rightarrow_a Q$ and $Q \Rightarrow_b P$. Then (i) $a = b = 1$ iff $min(a, b) = Cons(P, Q) = 1$,*
(ii) If either $a = 1$ or $b = 1$ then $min(a, b) \leq Cons(P, Q)$,
(iii) $min(a, b) = a \leq Cons(P, Q)$ iff $n \leq \frac{a-b}{a(b-1)}, a \neq 0, b \neq 1$,
(iv) $min(a, b) = a \geq Cons(P, Q)$ iff $n \geq \frac{a-b}{a(b-1)}, a \neq 0, b \neq 1$,
(v) $max(0, a + b - 1) \leq Cons(P, Q) \leq max(a, b) \leq s(a, b) = min(1, a + b)$.

The Cons function seems to be quite similar to a t-norm but not the same. So a closer look into the function is worthwhile.

We define a function $\star : [0, 1] \times [0, 1] \to [0, 1]$ as follows $\star(a, b) = \frac{a+b+nab}{n+2}$ where n is a non-negative integer.

Proposition 7. *(i) $0 \leq \star(a, b) \leq 1$,*
(ii) If $a \leq b$ then $\star(a, b) \leq \star(a, c)$,
(iii) $\star(a, b) = \star(b, a)$,
(iv) $\star(a, \star(b, c)) = \star(\star(a, b), c)$ iff $a = c$;
$\star(a, \star(b, c)) \leq \star(\star(a, b), c)$ iff $a \leq c$;
$\star(a, \star(b, c)) \geq \star(\star(a, b), c)$ iff $a \geq c$,

(v) $\star(a, 1) \geq a$, *equality occurring iff* $a = 1$,
(vi) $\star(a, 0) \leq a$, *equality occurring iff* $a = 0$,
(vii) $\star(a, b) = 1$ *iff* $a = b = 1$ *and* $\star(a, b) = 0$ *iff* $a = b = 0$,
(viii) $\star(a, a) = a$ *iff either* $a = 0$ *or* $a = 1$,

The consistency function $Cons$ gives a measure of similarity between two knowledges. It would be natural to define a measure of inconsistency or dissimilarity now. In [15] a notion of distance is available.

Definition 12. *If* $P \Rightarrow_a Q$ *and* $Q \Rightarrow_b P$ *then the distance function is denoted by* $\rho(P, Q)$ *and defined as* $\rho(P, Q) = \frac{2-(a+b)}{2}$.

Proposition 8. *The distance function* ρ *satisfies the conditions :*
(i) $0 \leq \rho(P, Q) \leq 1$
(ii) $\rho(P, P) = 0$
(iii) $\rho(P, Q) = \rho(Q, P)$
(iv) $\rho(P, R) \leq \rho(P, Q) + \rho(Q, R)$.
For proof the reader is referred to [15].

Definition 13. *We now define a measure of inconsistency by :*
$InCons(P, Q)$ = 1 - $Cons(P, Q)$

Proposition 9. *(i)* $0 \leq InCons(P, Q) \leq 1$,
(ii) $InCons(P, P) = 0$,
(iii) $InCons(P, Q) = InCons(Q, P)$,
(iv) $InCons(P, R) \leq InCons(P, Q) + InCons(Q, R)$ *for a fixed constant* n.

Proposition 9 shows that for any fixed n the inconsistency measure of knowledge is a metric. It is also a generalization of the distance function ρ in [15]; InCons reduces to ρ when $n = 0$. n is again a kind of constraint on the inconsistency measure - as n increases, the inconsistency increases too.

.

Definition 14. *We define consistency degree in the same way :* $Cons(\mathcal{C}_1, \mathcal{C}_2) = \frac{a+b+nab}{n+2}$ *where* $Dep(C_1, C_2) = a$ *i.e.,* $\mathcal{C}_1 \Rightarrow_a \mathcal{C}_2$ *and* $Dep(C_2, C_1) = b$ *i.e.,* $\mathcal{C}_2 \Rightarrow_b \mathcal{C}_1$.

Definition 15. *A measure of inconsistency for the case of covering in the same way is defined as follows :*
$InCons(P, Q)$ *= 1 -* $Cons(P, Q)$.

6 Towards a logic of consistency of knowledge

[11, 26] is the reference of our relevant papers for this section. We are now at the threshold of a logic of consistency (of knowledge). Along with the usual propositional connectives the language shall contain two binary predicates, '$Cons$' and

'Dep' for consistency and dependency respectively. At least the following features of this logic are present.

(i) $0 \leq Cons(P,Q) \leq 1$,
(ii) $Cons(P,P) = 1$,
(iii) $Cons(P,Q) = Cons(Q,P)$,
(iv) $Cons(P,Q) = 0$ iff $Dep(P,Q) = 0$ and $Dep(Q,P) = 0$
and $Cons(P,Q) = 1$ iff $Dep(P,Q) = 1$ and $Dep(Q,P) = 1$
In case P,Q,R partitions we also get
(v) $Cons(P,Q)$ and $Cons(Q,R)$ imply $Cons(P,R)$.

(i) shows that the logic is many-valued; (ii) and (iii) are natural expectations; (iv) conforms to assumptions 1 and 2 (section 5); (v) shows transitivity the predicate Cons in the special case of partitions.

7 Concluding remarks

We have considered almost all the existing lower / upper approximations in rough set literature and there properties are given in tabular form showing briefly their behavior patterns.

We draw the pictures of partially ordered sets and implication lattices which show independent equivalence classes and also the node at the tail implies that at the head.

Any of the inclusion gives rise to a rough Modus Ponens rule [6] and a corresponding rough logic [6]. The underlying modal logical systems of various rough logics are also immediately visible from the tables.

In future we wish to work on the topological and logical aspects of covering based approaches of Rough Set Theory.

We proposed Consistency-Degree between two knowledges and we want to develop an interesting many valued logic of dependency and consistency of knowledges.

References

1. Banerjee,Mohua., and Mihir K. Chakraborty : Rough Sets Through Algebraic Logic. Fundam. Inform. 28(3-4): 211-221 (1996).
2. Banerjee,Mohua., and Mihir K. Chakraborty : Algebras from Rough Sets. Rough-Neural Computing - Techniques for Computing with Words. Sankar K. Pal, Lech Polkowski, Andrez Skowron (Eds.). © Springer - Verlag Berlin Heidelberg New York, 2004.
3. Banerjee, Mohua and Yiyu Yao: A Categorial Basis for Granular Computing. RSFDGrC 2007: 427-434.

4. Bianucci,D.; Cattaneo,G. and Ciucci,D. : Entropies and co-entropies of coverings with application to incomplete information systems, Fundamenta Informaticae 75 (2007), 77-105.
5. Bonikowski, Z.: A Certain Copnception of the Calculus of Rough Sets. Notre Dame J. Formal Logic 33, 412421 (1992).
6. Bunder, Martin W; Banerjee, M and Chakraborty, M.K. : Some Rough Consequence Logics and their Interrelations, J.F.Peters and A. Skowron(Eds.) : Transactions on Rough Sets VIII, LNCS 5084, pp. 1-20, 2008. © Springer-Verlag Berlin Heidelberg 2008.
7. Cattaneo, Gianpiero : Foundational and Mathematical Invcestigation of Roughness Theory. Preprint.
8. Cattaneo, G., Ciucci,D. 2009. : Lattices with interior and closure operators and abstract approximation spaces. In; Peters,J., Skowron,A.(Eds.). Foundations of Rough Sets X. Springer, pp. 67-116.
9. Cattanio,G. and Cucci,D. : Lattice Properties of Preclusive and Classical Rough Sets. Personal Collection.
10. Chakraborty,M.K. and Banerjee,M. : Rough dialogue and implication lattices, Fundamenta Informaticae, 75(1-4) , 2007, 123-139.
11. Chakraborty,M.K. and Samanta,P. : (2007) Consistency-Degree Between Knowledges. M.Kryszkiewicz et al. (Eds.): RSEISP 2007, LNAI 4585, pp. 133-141. © Springer-Verlag Berlin Heidelberg 2007.
12. Klir,G.J. and Yuan,B. : (1997) FUZZY SETS AND FUZZY LOGIC : Theory and Applications, Prentice-Hall of India.
13. Li, Tong-Jun : Rough Approximation Operators in Covering Approximation Spaces. S. Greco et al. (Eds.): RSCTC 2006, LNAI 4259, PP. 174-182, 2006. © Springer-Verlag Berlin Heidelberg 2006.
14. Liu,Jiao and Liao,Zuhua : The sixth type of covering-based rough sets. GrC 2008 IEEE International Conference on Granular Computing 2008, pp. 438-441.
15. Novotný,M. and Pawlak,Z. : (1988) Partial Dependency of Attributes. Bull. Polish Acad. of Sci., math., 36, pp.453-458.
16. Pal, Sankar K.; Polkowski, Lech; Skowron, Andrej. : Rough - Nural Computing : Sankar K. Pal, Lech Polkowski, Andrej Skowron(Eds.) Springer-Verlag 2004.
17. Pawlak,Z. : Rough Sets. International Journal of Computer and Information Sciences, Vol. 11, No. 5, pp.341-356, 1982.
18. Pawlak,Z. : (1985) On Rough Dependency of Attributes in Information System. Bull. Polish Acad. of Sci., Math., 33, pp.551-559.
19. Pawlak,Z. : (1991) ROUGH SETS - Theoritical Aspects of Reasoning About Data. Kluwer Academic Publisher.
20. Pagliani, Piero : "Pre-topologies and Dynamic Spaces": RSFDGrC 2003, hongqing, R.P. China, 2003, 146-155. Extended Vertion in Fundamenta Informatica, 59(2-3), 2004, 221-239.
21. Pagliani,Piero., and Mihir K. Chakraborty : A geometry of Approximation - Rough Set theory: Logic, Algebra and Topology of Conceptual Patterns. © Springer Science + Business Media B.V. 2008.
22. Pomykala,J.A. Approximation operations in approximation space , Bulletin of the Polish Academy of Sciences: Mathematics 35 (1987) 653-662.
23. Pomykala,J.A. : Approximation, Similarity and Rough Constructions. ILLC Prepublication Series for Computation and Complexity Theory CT-93-07, University of Amsterdam(1993).
24. Pomykala, J., & Pomykala, J. A. : "The Stone Algebra of Rough Sets". Bull. Polish Acad. Sci, Math., 36(7-8), 1988, 495-508.

25. Qin,Keyun; Gao,Yan and Pei,Zheng : On Covering Rough Sets. J.T. Yao et al. (Eds.): RSKT 2007, LNAI 4481, pp. 3441, 2007. ©Springer-Verlag Berlin Heidelberg 2007.
26. Samanta,P. and Chakraborty,M.K. : On Extension of Dependency and Consistency Degrees of Two Knowledges Represented by Covering. J.F. Peters et al. (Eds.): Transactions on Rough Sets IX, LNCS 5390, pp. 351-364, 2008. © Springer-Verlag Berlin Heidelberg 2008.
27. Samanta,P. and Chakraborty,M.K. : Covering Based Approaches to Rough Sets and Implication Lattices. H. Sakai et al. (Eds.): RSFDGrC 2009, LNAI 5908, pp. 127134, 2009. © Springer-Verlag Berlin Heidelberg 2009.
28. Samanta,P. and Chakraborty,M.K. : Generalized Rough Sets and Implication Lattices. J.F. Peters et al. (Eds.): Transactions on Rough Sets XIV, LNCS 6600, pp. 183201, 2011. © Springer-Verlag Berlin Heidelberg 2011.
29. Slezak, Dominik and Wasilewski, Piotr : Granular Sets Foundations and Case Study of Tolerance Spaces. A. An et al. (Eds.): RSFDGrC 2007, LNAI 4482, pp. 435442, 2007. ©Springer-Verlag Berlin Heidelberg 2007.
30. Yao,Y.Y. : Constructive and algebraic methods of the theory of rough sets. Journal of Information Sciences 109 (1998) 21-47.
31. Yao,Y.Y. : On generalizing rough set theory. Rough set, Fuzzy sets, Data Mining and Granular Computing. Proceedings of the 9th International Conference (RSFDGrC 2003), LNAI 2639, pp. 44-51, 2003.
32. Yun, Ziqiu and Ge, Xun : Some notes on covering-based rough sets. Preprint.
33. Zhu,William : Topological approaches to covering rough sets. ScienceDirect, Information Sciences 177 (2007) 14991508.
34. Zhu,William : Relationship between generalized rough sets based on binary relation and covering. Information Sciences 179 (2009) 210-225.
35. Zhu,William and Wang,Fei-Yue : Relationship among Three Types of Covering Rough Sets. Proc. IEEE Int'l Conf. Grannuler Computing (GrC ' 06), pp. 43-48, May 2006.

Pulak Samanta
Department of Mathematics,
Katwa College, West Bengal, India.
pulak_samanta06@yahoo.co.in

Journey and Joy of Mathematics

Satyabachi Sar

No one can ascertain that auspicious moment when mathematics started its journey. But it is certain that it started its journey slowly and silently in the prehistoric time. Great 'epic poet' Homer rightly remarked about mathematics that

> 'Small at her birth but rising every hour
> She stalks the earth and shakes the world around.'

Starting its journey in a humble way mathematics, now, occupies a predominant position in science, society and every where. With passage of time mathematics has changed its perspective and priority. It has been described in different ways by various persons at different times. Evolving from counting, calculation, measurement mathematics has been expanded to such an extent that any thing and every thing can be brought under the purview of mathematics. Many branches of mathematics have emerged to meet the necessity of time and society. At present, mathematics is being used throughout the world as an essential tool in many fields of utility, which include natural science, engineering, technology, medicine, social sciences. Mathematics, too, may be used as a source of joy in many respects, if we look to it properly in search of beauty, fun, thrill, excitement concealed in it.

Mathematics has come down to the present as an outgrowth of thought that originally centered in the concept of number and form which were carried along arithmetic, algebra and geometry. In the initial stage mathematics arose as a part of every day life.

The first conception of number and form started in the Old Stone Age, the Paleolithic. When the old stone age gave way to the New Stone Age, the Neolithic, transition in the way of livelihood of man occurred from the mere gathering food to its actual production, from hunting and fishing to agriculture. This happened perhaps ten thousand years ago. Numerical terms came slowly into use. Their first occurrence was 'qualitative rather than quantitative', making a distinction only between one and two and many. Neolithic man also developed a feeling for geometrical pattern.

During the period of fifth, fourth and third millennia B.C. more technically advanced society emerged. Ancient civilizations like Babylonian, Egyptian, Indian and Chinese civilizations came to flourish along the banks of the rivers of the Tigris and the Euphrates, the Nile, the Indus later the Ganges, the Huag He later the Yanzi respectively. Mathematics grew there as a practical science. It may been inferred that the essence of general arithmetical and algebraic nature evolved in these civilizations was very much alike. Historians of mathematics have given account of evidences of Babylonian mathematics (c. 2000 B.C.), Egyptian mathematics (c. 1850 B.C.), Indian mathematics (c. 1800 B.C.), Chinese mathematics (c. 1200 B.C.). But it is to be remembered that more complex mathematics did not appear when these civilizations began to use arithmetic, algebra and geometry for different purposes such as computation of calendar, taxation and other financial calculations, survey, building and construction, astronomical observations etc.

Though the Greeks inherited the Egyptian and Babylonian basis of mathematics when transition of intellectual leadership from the valleys of the Nile, Tigris and Euphrates rivers was made to the shores of Mediterranean, independent development of mathematics took place. The systematic study of mathematics in its own right began here. Empiricism gave way to deductive reasoning. Necessity of proof was felt in an earnest way for proper evolution of mathematics. It is to be noted that mathematical proof is fundamentally a matter of rigour.

Really, 'without the strictest deductive proof from admitted assumptions, explicitly stated as such, mathematics does not exist'. But at the same time we may not ignore the fact that intuition, experiment, induction and plane guessing are important elements in mathematical invention.

With the recognition of strict deductive reasoning mathematics began to emerge around 600 B.C. The emergence was complete when 'abstraction' came to exist in mathematics. It is rightly remarked that 'abstracting of common experience is one of the principal source of utility and the secret of its scientific power.'

It is to be observed that 'mathematics as an organized, independent and reasoned disciple did not exist before the classical Greeks of the period from 600 to 300 B.C.' But the major Greek progress in mathematics may be estimated from 300 B.C. to 200 A.D.

It is true that story of great mathematicians who were the representatives of their time in mathematics is an essential component in describing development of mathematics. Also their story guides us to gather knowledge about how a mathematician thinks, how his imagination, as well as, his reason leads him to new aspects of truth. But for the paucity of space we restrict ourselves only to mention the names of some great mathematicians with their celebrated contributions.

Contribution of Thales (c. 626 - c. 545 B.C.), Pythagoras (c. 584 - c. 495 B.C.), Zeno (c. 495 - c. 435 B.C.), Hippocrates (c. 460 - c. 370 B.C.), Euclid (c. 330 - c. 275 B.C.), Archimedes (c. 626 - c. 545 B.C.), Ptolemy (c. 85 - c. 165 A.D.), Diophantus (around 250 A.D.), Pappas (around 320 A.D.) and others added to great progress of the Greek mathematics, which mainly consisted of geometry, number theory and algebra. The Greek gave 'three famous problems of antiquity', namely, the trisection of an angle, the duplication of cube and the quadrature of the circle, which created great sensation among the future mathematicians and let them to resolve the problems.

Mathematics also flourished in India and Islamic world. Discovery of zero and place value system of numeration in India and contribution of mathematicians like Aryabhata (476-550 A.D.), Brahmagupta (598-670 A. D.), Al-Khwarizmi (c. 780-850), Omor Khayyam (c. 1050-1130), Bhaskara II (1114-1185) brought bright light to illuminate mathematics.

Now, into the two main streams of number and form flowed many tributaries. Counting by natural numers 1,2,3,4,... gave birth to 'discreteness'. Then came the concept of 'continuity'.

Specially with the admirable effort of Fibonacci (c. 1180 - 1250), the knowledge of Islamic mathematics mainly borrowed from India, along with the knowledge of the Greek mathematics was brought to Europe. Mathematics began to flourish there. Major progress of mathematics was made in the beginning of the sixteenth century and stimulation in mathematical activities continued unabated.

In addition to number, form, discreteness and continuity, a fifth stream emerged. The fifth stream was analytic geometry, which has been of great importance in the history of mathematics.

Many celebrated mathematicians like Cardano (1501-1576), Bombelli (1526-c. 1573), Viete (1540-1603), Napier (1550-1617), Gallilei (1564-1642), Kepler (1571-1630), Descartes (1596-1650), Fermat (1601-1665), Pascal (1623-1662), Hygen (1629-1695), Newton (1642-1727), Leibniz (1646-1716), etc. appeared in the scene. They were instrumental to give birth to new branches of mathematics and their immortal contribution to enrich different branches of mathematics made them unparallel in the history in their own fields of investigation.

Geometry became analytic in 1637 with publication of Descartes' masterpiece. In the seventeenth century the three main streams of number, form and continuity united and led to generate calculus in an organised way and mathematical analysis, in general. Newton and Leibniz, building on the works of many mathematicians, established calculus on a strong footing. Dynamics of Gallilei and Newton began to be common property of all creative mathematicians.

In the eighteenth century Euler (1707-1783), Lagrange (1736-1813), Legendre (1752-1833), Cauchy (1759-1857), Fourier ((1768-1830), Gauss (1777-1855) and others contributed in a tremendous way to lift mathematics to an extraordinary level.

Euler invented two new branches of mathematics, namely, calculus of variation and differential geometry in addition to work in a wide range of mathematical branches. He also pushed forward research in number theory which was effectively begun by Fermat.

It is necessary to mention here that symbols and notations in mathematics play a vital role in advancement of mathematics and understanding concepts of mathematics. Really, 'symbolic notation is half of mathematics', as mentioned by Bertrand Russel (1872-1970). For enrichment of mathematical symbols and notations, contribution of many mathematicians, specially of Viete and Euler, should respectfully be remembered.

Towards the end of the eighteenth century Lagrange began rigorous theory of functions and theory of mechanics.

1801 marked the beginning of a new era with the publication of Gauss' masterpiece. Also his work in differential geometry was revolutionary. His contribution to astronomy and mechanics was startling. Really, in the nineteenth century the great river of mathematics 'burst its banks, deluging wilderness where no mathematics had flourished and making them fruitful'. Cauchy began marvelous treatment on analysis and initiated study of the theory of functions of complex variable. Weierstrass (1815-1897) and Riemann (1826-1866) continued this work. Fourier's work on heat proved to be of fundamental importance.

Invention of non-Euclidean geometry by Lobachevsky (1793-1856) and Janos Bolyai (1802-1860) led to characterization of geometry by Riemann. Cayley (1821-1895) advanced algebraic geometry and his work on matrices and linear algebra was complemented by Hamilton (1788-1856) and Grassmann (1809-1877). Inventing quaternions Hamilton opened the door of algebra and algebra became algebras as through the invention of non-Euclidean systems, geometry became geometries. Galois (1811-1832) introduced group concept which gave new direction in mathematical research.

The end of the nineteenth century witnessed the spectacular discovery of Cantor (1845-1918). He introduced theory of sets and theory of transfinite numbers. His analysis of the concept of number added to the major work of Dedekind (1831-1916) and Weierstrass on irrational numbers.

Journey of mathematics has made more progress with progress of time. It would be a fascinating story if one tries to describe further development of mathematics, which has been possible by hard work, strong will, dedication, sacrifice, perseverance of many mathematicians. But I want top stop here just mentioning that the persons, who were instrumental in advancing the mathematical ideas, have done it not only for the benefit of science and society but also for the feeling of ecstatic joy derived from the theory, structure, pattern, symmetry, utility of mathematics. At the same time we should understand and realise the fact that a lot of difficulty, struggle, obstacles have been faced by stalwarts of mathematics, though gleefully, to invent different components of modern mathematics such as non-Euclidean geometries, groups, general relativity, theoretical physics, set theory, non-standard logic, fuzzy set theory, rough set theory, etc.

Regarding the journey of mathematics through ages my humble submission is that during its description a few elements of mathematics or a few incidents which may be considered by other to be quite significant, may have been left out, as the journey is so long, elaborate and stretched that all the elements and events cannot be accommodated in a small article like this in which only a brief account is given and also as priority of choice varies from one to another depending on his or her point of view.

Leaving journey of mathematics aside we now turn to joy of mathematics. First of all let us see what does joy mean. Joy is a feeling of exultant happiness. When such a situation or a condition or an occasion arises that our heart is filled with ecstatic satisfaction or happiness we seem to swim in an ocean of joy. Indeed, joy may be derived from various sources. These sources include a unique piece of music, a melodious song, a spectacular sculpture, a marvelous painting, a panoramic view of nature, so on. Mathematics may be a source of joy if we look to its beautiful formula, equation, pattern, arrangement etc. with an eye of sensitiveness or a mind of discovery. Also charm or challenge or fun of mathematics may often emerge as a source of joy. Sometime a glimpse to a novel concept or to an intriguing idea or to a historical anecdote of mathematics not only provides us joy but also arouses our curiosity to seek, to search, to explore more delight and magnificence from it.

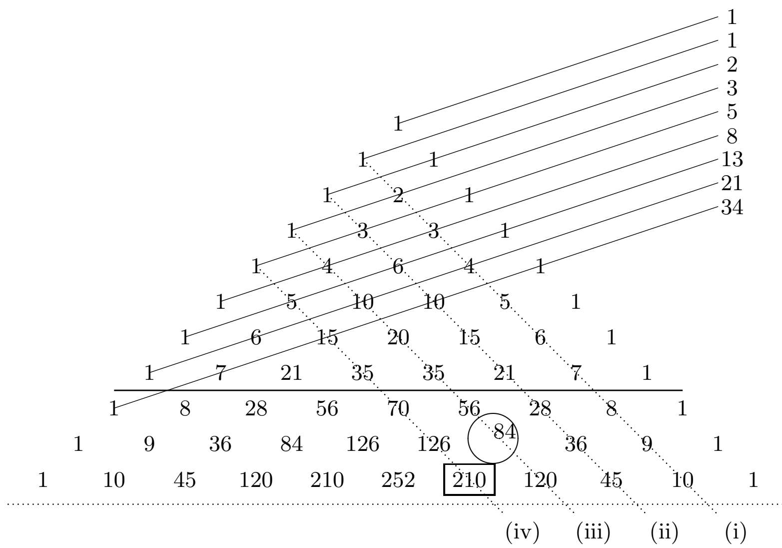
Here we state a few mathematical objects from which a mathematical mind can easily obtain immense joy and unfathomable beauty.

We start with so called Pascal triangle, which is an infinite numerical triangular array of natural numbers, obtained as co-efficients of binomial expansions of indices successively increasing one by one staring from zero. I intentionally used so called Pascal triangle, not 'Pascal triangle', because of the fact that the mathematical concept was known to India, Persia, China and Italy centuries before French mathematician B. Pascal (1623-1662).

Before going to narrate the beauty of this numerical arrangement, let us give its historical account. The numerical table appeared in 'Traité du triangle arithmétique' (Treatise on arithmetic triangle) written by Pascal in 1653 and was published in 1665 posthumously. Hogben (P.323) claimed that Persian poet and mathematician Omar Khayyam (c. 1052 - 1130) described this numerical series and it was figured in the work of Chinese algebraist Chu-Shih Chien (fl. 1280-1303). Boyer (P. 231), too, was in the same opinion about Chu-Shih Chien. According to Srinivasienger (P. 27-28) these numbers were described by Pingala in third century B.C. in his 'Chhandas-Sutra', but the rule for determining the numbers, described by him was difficult. It was explained by commentator Halayudha in 1000 A.D. and a numerical diagram called 'Meru Prastara' was formulated. The rule was nothing but the formula $n+1_{C_r} = n_{C_r} + n_{C_{r-1}}$.

Really, it would be a matter of great joy if with the help of this numerical triangle one can find interconnection of different mathematical ideas discovered by different persons at different time in different places. Here we will see that binomial formula discovered by Newton, terms of the sequence named after Fibonacci, figurate numbers of Greek mathematics, probability discovered by Pascal have direct connection with this numerical table prevalent in the name of Pascal.

Let us first describe the table below:



Here we observe that dotted diagonal (i) represents natural numbers, dotted diagonal (ii) represents triangular numbers, dotted diagonal (iii) represents tetrahedral numbers and dotted diagonal (iv) represents 4-space tetrahedral numbers.

We know that Newton's binomial formula is given by
$(a+b)^n = a^n + n_{C_1} a^{n-1} b + n_{C_2} a^{n-2} b^2 + ... + n_{C_r} a^{n-r} b^r + ... + n_{C_n} b^n$.
It is to be noted that the co-efficient of binomial expansion with index n are depicted in the $(n+1)$-th row of table. One can easily verify that the 8th row, which has been underlined here, contains the co-efficient of $(a+b)^7$.

Next we consider Fibonacci sequence $\{F_n\}$, which is defined by the recurrence formula
$F_n = F_{n-1} + F_{n-2},\ n \geq 3,\ F_1 = F_2 = 1.$
So the terms of Fibonacci sequence are 1, 1, 2, 3, 5, 8, 13, 21, 34,... , which are also known as Fibonacci numbers.

It is interesting to note that Fibonacci numbers can be obtained from the above numerical table by summing up the numbers in the diagonals as indicated in the table. One can observed that the numbers are obtained as 1, 1, 2(=1+1), 3(=1+2), 5(=1+3+1), 8(=1+4+3), 13(=1+5+6+1), 21(=1+6+10+4),
34(=1+7+15+10+1), and so on.

It is known that the nth triangular (or (2-D)) number, nth tetrahedral (or (3-D)) number, nth 4-space tetrahedral (or (4-D)) number, nth 5-space tetrahedral (or (5-D)) number,... are obtained respectively as the sum of first n naturals (or (1-D)) numbers,

first n terms of (2-D) numbers, first n numbers of (3-D) numbers, first n terms of (3-D) numbers,..., given by
$1+2+3+...+n=\frac{1}{2!}n(n+1)$
$1+3+6+...+\frac{1}{2}n(n+1)=\frac{1}{3!}n(n+1)(n+2)$
$1+4+10+...+\frac{1}{6}n(n+1)(n+2)=\frac{1}{4!}n(n+1)(n+2)(n+3)$
$1+5+15+...+\frac{1}{24}n(n+1)(n+2)(n+3)=\frac{1}{5!}n(n+1)(n+2)(n+3)(n+4)$
and so on.

It is quite pleasing surprise to observe that (1-D), (2-D), (3-D), (4-D), (5-D),... are numbers appear in the table as shown in the dotted diagonals.

It is very interesting to state that the sum of the numbers down any dotted diagonal to a particular one can easily be located in the so called Pascal Triangle. The sum would be equal to the number which is one place left to the number that is directly below the particular number in the next row. For example to locate the sum of triangular numbers starting from 1 to 28, we find the triangular number in the next row, which happens to be 36 and encircle the number 1-place left to 36 in the row and get the sum as 84. In the similar way, if we want to find the sum of tetrahedral numbers starting from 1 to 84, we go to 120, the next tetrahedral number in the next row and go to one place left to find the sum [Here we put a rectangle around the sum]. The method holds good for all the numbers in all the dotted diagonals i.e., for (1-D), (2-D), (3-D), (4-D), (5-D),... numbers.

More interestingly, one can link probability to this numerical triangle in various ways, which I am not going to show for shortage of space.

There are many more mathematical objects from which one can gather very interesting and astonishing interconnections among various mathematical thoughts that can be treated as source of joy. But I like to stop here just mentioning a thrilling news which may provide immense joy to the lovers of mathematics. In 1970, Nicaragua, a country situated in Central America, issued ten postage stamps bearing the ten mathematical formulas 'that changed the face of the world'. 'Isn't it admirable that a country so respects mathematics that it devotes a postage-stamp series to a set of abstract equations ?' But basis of selection of formulas to be elevated to high a status is not known, though one thing is known that $e^{\pi i}+1=0$, is not included. But according to many it is 'a celebrated equality' in which there are 'the five most significant numbers (as well as the most important relation and most important operation) in all mathematics'. Not only it has been termed by Richard Fineman as 'the most beautiful equation in mathematics', but also Turnbull, in his book 'The Great Mathematician', remarked 'Was it not Felix Klein who remarked that all analysis was centre here ? Every symbol has its history — the principal whole number 0 and 1; the chief mathematical relation + and =, π the discovery of Hippocrates; i the sign for the impossible square root of minus one; e the base of Napierian logarithm'.

References

1. Bell, E.T. : *The Development of Mathematics*; McGrow-Hill Book Company, New York(1945).
2. Boyer,C.B. : *A History of Mathematics*, John Wiely and Sons, Singapore(1989).

3. Cajori, F. : *A History of Mathematics*, The Macmillan Company, New York(1953).
4. Gulberg,J. : *Mathematics from birth of numbers*, W.W. Norton & company, New York(1997).
5. Hogben, L. : *Mathematics for the Million*, George Allen and Uncoin Limited, London(1960).
6. O'Connor, J.J. and Robertson E.F. : *An overview of the history of mathematics*, http://www-history.mcs.st-and.ac.uk/HistTopics (2011).
7. Newman, James R. (ed.) : *The World of Mathematics, Volume 1*, Dover Publications, New York(2000).
8. Pappas, T. : *More Joy of Mathematics*, Wide World Publishing, Tetra, San Carlos(1998).
9. Pickover, C.A. : *Wonders of Numbers*, Oxford University Press, New York(2001).
10. Sar, Satyabachi : *Jana-Ajanay Ganitik Dharona, Chinha O Sanket*, (in Bengali), Jnan-Bichitra, Agartala, India, 2010.
11. Smith, D.E. : *History of Mathematics (Vol. I and II)*, Dover Publications, New York(1958).
12. Srinivasienger, C.N. : *History of Ancient Indian Mathematics*, World Press Private Ltd, Calcutta(1988).
13. Struik, D.J. : *A Concise History of Mathematics*, Dover Publications, New York(1986).

[I feel immensely happy to be requested to contribute an article to celebrate the birthday of Prof. Mihir K. Chakraborty. Prof. Chakraborty happens to be not only a very close friend of mine but also a sensational personality in many fields, specially as a stalwart in the world in his field of research. More over, he is a very good human being who always tries to help his students, friends and others. I shall be failing in my duty if I do not mention that Prof. Chakraborty has given me enough inspiration and support for pursuing mathematical works in the field of pedagogy and history of mathematics. He has been instrumental in providing me encouragement in writing a lot of articles, specially in Bengali, on popular, semi popular, serious mathematics in science magazines which have culminated in the publication of a few books. I sincerely wish him a very long active life with good and sound health providing fruitful contribution to mathematics and society, which Prof. Chakraborty has been doing so far.]

Satyabachi Sar
S. N. Bose Birth Centenary Research Fellow,
Calcutta Mathematical Society, India,
and
Retired Faculty, Department of Mathematics,
M.B.B. College, Agartala, India.
`satyabachisar@gmail.com`

Mathematics and music: An aesthetic detour into logic

Sundar Sarukkai

Mihirda is not only a creative mathematician and logician but he also brings his inquisitive reflection to bear on the fields of art. The best homage I can pay to him is by engaging with the question of mathematics and music, a topic on which he himself has written in his Bengali text.

Let me begin with a question on the logic of fiction. In one of the seminars at the Department of Comparative Literature, Jadavpur University, where Mihirda was also present, I spoke on the relation between fictional and scientific discourse. There was one point which interested him and I want to begin with that particular issue.

Scientific discourse is imaginary and 'fictional' even though its primary purport is to describe a real world. The use of mathematics in this discourse makes this relation between reality and science far more complex. Either we accept mathematical entities as somehow being real and thus having an innate capacity to capture the reality of our universe or we deny these entities any status of reality. The former is basically the position of Platonism but Platonic reality is so esoteric that the relation between physical reality and Platonic objects remains a serious problem. The latter position is closer to viewing mathematical entities as fictional entities and the challenge here is to understand why then do such entities matter so much to the description of the real world.

In contrast, we have the fictional discourse as exemplified in literature. Although the fictional discourse presents itself primarily as something unreal there is a significant space of the real 'present' in it. One of the best ways to understand societies and people is to read novels. Fiction captures the real in very important ways. The modifications it makes to real events are not significant enough to detract from the enduring presence of the real in fiction.

This claim can perhaps be best illustrated by considering two genres of fiction that by definition cannot have anything to do with the real: fantasy and science fiction. If we analyse the characters, the events, the language and the psychological processes in these stories, we can notice a strong influence of human reality even in these genres. Fiction does indeed have a serious problem in escaping the influence of the real even though it presents itself as something unreal and untrue.

This paradoxical reversal of the nature of scientific and fictional discourse suggests a simple but profound philosophical truth: a discourse of the real has to perforce depend on the metaphysics of the unreal. Mathematics plays an important role in making this relation possible. In my talk at Jadavpur I had suggested that there is a logic that seems to be at the foundation of this problem.

The logic is as follows: science is “true and *possibly* false” whereas literature (fiction) is “false and *possibly* true”. My intuitive idea was that a logic of science and literature would have to be a mix of logics, including modal logic.

I believe that this formulation captures a simple truth about the nature of science and fiction. As extensive literature in the philosophy of science points out, scientific knowledge is fundamentally fallible, that is, open to revision. But until and unless there are reasons for revision, knowledge as of that moment is taken to be true. I would like to formulate the logic of fallibility as the logic of True and Possibly False. Similarly, fictional knowledge (knowledge that arises from reading fiction) is fundamentally false and does not want to be taken as true. But when faced with the challenge of explaining how close to the real the fictional discourse is, it can hide behind the excuse that fiction can possibly be true. In claiming this, fiction can continue to hold on to its claim that it is essentially about falsity. This logic of the fictional can thus be summarized as False and Possibly True.

It was Mihirda who immediately caught on this point. I did plan to discuss it in greater detail with him but had no opportunity to do so. Unfortunately, I am not going to discuss this issue in any detail here also but am only using it to enter into a discussion on the relation between aesthetics and logic.

One might find it odd to bring the question of logic to bear on fiction, since fiction prides itself on being consciously illogical. But fiction is fundamentally enmeshed in this type of a logical conundrum. The first point is about the inherent modality present in this logic. I believe that it is possible to locate the aesthetic within this modality. In other words, the logic of the aesthetic is to be found in the idea of the ‘possibly true’ (in the case of fiction) and ‘possibly false’ (in the case of science).

What do I mean by saying that the logic of the aesthetic is to be found in these modal ideas? Science is fundamentally indebted to the aesthetic imagination but even when scholars note this point they often place the logical against the aesthetical. That is, after recognizing that scientific discoveries may have been catalysed by an aesthetic impulse, they place the real significance of that within the formal scientific structures such as theory. (This should remind us of the well-known distinction, coined by Reichenbach, between contexts of discovery and contexts of justification in science. The aesthetic imagination works in the former realm and the logical in the latter.)

In contrast, my argument is that one way of understanding the essential role of the aesthetic imagination in science is by recognizing that aesthetics arises in the process of acknowledging that there is something possibly false. The exploration of the possibly false in true statements is the defining moment of creativity and the creation of new science. If scientists accepted the truthfulness of scientific statements then how is new science possible? New ideas in science are catalyzed by the belief that the true statements of science are possibly wrong and the attempt to prove them wrong generates new ideas which are then justified to be true.

Similarly, if fiction is accepted as only being false, what then is the genesis for creating new fiction, for continuously creating the false? That new creation is possible only through the recognition, the constant recognition, that fiction is possibly true. The origin of aesthetics in both these cases arises from this impulse mediated through the logic of possibility.

Thus, both science and art are enmeshed within the ideas of true and false, as well as with possibility. Is this true of art in general and music in particular?

In contrast to common belief about art, there are strong philosophical traditions, whether in the Indian aesthetic tradition or the phenomenological tradition in Europe exemplified by Heidegger and Gadamer, which argue that art is primarily about truth. While much has been written on this in the context of visual art, what exactly does it mean in the context of music?

First of all, music does not express or even attempt to capture propositional truths. But the idea of truth is far deeper than propositional truths. Logic, dealing as it does with propositions, might therefore seem to be far removed from music. However, there is another way of understanding the musician's idea of logic within music, even though they may not phrase it as such. If we are sympathetic to the understanding of logic as an activity that primarily has to do with sentences, in a manner similar to Davidson, then the idea of logic within music is apparent. Logic is a series of relations between sentences and thus the idea of logic is fundamentally in the ideas of succession and intertwinement between phrases and sentences.

Music is also basically about succession and intertwinement between musical phrases and 'sentences'. If notes in music are seen as 'words' then one can see how phrases of notes and combinations of notes function as larger units whose intertwining and succession lead to the sense of logic within music. We should remember here that the most fundamental formulation of music always comes back to the idea of succession. Music, as also understood in the phenomenological tradition, is an exemplar of time consciousness where time is a series of sequences. This immediately suggests that the logic of music is perhaps more tuned to temporal logic but at this point I hesitate to make this claim.

Yet another illustration of the logic within music is through its association with mathematics. All classical traditions of music perforce engage with the idea of some intricate mathematics. Perhaps this is the reason why they are called as classical music because of a theory of music, often including intricate computation, which forms the base of this music. Music and mathematics come together through the notion of time. Whether in the idea of beats, taan, taal, rhythm or even melody, time units play a fundamental role in the aesthetic appreciation of music. For example, in Hindusthani music, the concept of laya is among the most important one. As the well-known musician Dinkar Kaikini points out, 'laya is time' but it is an 'abstract continuum'. The importance of laya is captured in his comment that 'when you divide laya aesthetically, it becomes music.'

Counting of time and keeping time is essential to any production of music. Time is kept through intricate counting mechanisms which have a strong mathematical basis. In Hindustani music, for example, this relationship with time is importantly not just for the tabla player but also for the singer. The process of computing that is essential to good percussion musicians like the table and mrudangam players is a wonderful illustration of how mathematics is converted to a domain of feeling. When these musicians play their instruments, they are not computing before playing. They can feel the rhythm, can feel when a beat is skipped or can feel how to create a complex mathematical pattern in the middle of a concert. (I believe that this sense of 'feel'/contact/touch with mathematics is experienced by many mathematicians. Cantor's claim that one can 'see' sets as clearly as other objects might perhaps seem absurd to some but if we replace 'see' with 'feel', I have a *feeling* that many more mathematicians would recognize that as a common process.)

This applicability of mathematics in music is as mysterious as the applicability of mathematics in the sciences. In the latter case, scientists have often wondered why mathematics is so effective in describing a world (the physical universe) that is so different from the world of mathematics (the abstract or Platonic world). The success of the applicability of mathematics in the sciences is illustrated in the success of theory among which one can point to the power of unification of diverse phenomena and the predictive capacity through mathematical models.

The mysteriousness of the applicability of mathematics in music is fundamentally about the creation of the aesthetical from the mathematical. As much as there is a problem in discovering truths about the physical world through the use of mathematical techniques, there is an equally mysterious problem in this relation between our musical sensibilities and mathematical and computational complexity. Why should the mathematical imagination not only matter to musical sensibilities but also create the musical experience?

There are some essential points of difference, for example, in the quantity and the types of mathematical entities and structures, between applicability of mathematics in music and in science. However, the basic notion of applicability is the same and similar philosophical questions arise in both these cases. Moreover, the presence of mathematics in music alerts us to the possibility that the use of mathematics need not be conscious but can be something 'felt' through years of training and experience. The most interesting question is this: How does this form of mathematical applicability lead to an aesthetic experience? Mathematical applicability in science leads to knowledge and truth but in the case of music it leads to profound experiences. And we should not forget that this experience of music is fundamentally auditory so somehow there is a conversion or expression of the mathematical into the concrete auditory sensation and through this process an aesthetic experience is formed. (In contrast, in science, the aesthetics associated with mathematics is fundamentally a visual process associated with 'seeing' mathematical structures in the way they are written and calculated.)

The use of mathematical and computational techniques in music must alert us to a fundamental concept that underlies both science and art in general and music in particular. The invocation of mathematics (as against logic) in an explicit manner is an indication of the importance of the idea of truth in both these activities. Truths of nature cohabit with truths of the mathematical domain for science. Truths of aesthetic judgement cohabit with truths of the mathematical for music. Therefore, it should not surprise us that for both science and music, the question of logic is never prominently foregrounded. While there are underlying logical structures (including logic through mathematics) in both science and music, the preoccupations of logic are never explicitly invoked in science. Thus, while there is an unreasonable dependence on mathematics in science, there is no such dependence on logical principles, theorems and results in the discourse of science. Similarly in the case of music – the very mention of logic in the context of music generates heated disagreement!

But it is important to note that both science and music engage very deeply with the idea of truth. Scientific discourse is filled with references to truths of the world, truths of universal structures like laws, truths of mathematics and so on. There is also a long and enduring relation between truth and beauty which scientists have drawn freely upon. This is well captured in the title of the book 'Truth and Beauty' by S. Chandrasekhar, the Noble Prize winner in physics. In many studies of aesthetics in science, one often stumbles upon this repeated invocation of the relation between truth and beauty. Not only do science and music depend on the notion of truth but they also share another interesting trait. Both of them have a problematical relationship to the idea of meaning. While the question of meaning is the bugbear of language, science – even though it is primarily a discourse – rarely engages with the problems of linguistic meaning. Similarly, music – and art in general – finds the repeated questioning of the meaning of an artwork itself as meaningless. Nothing irritates an artist, including musicians, than asking what does your performance *mean*?

What could be the notion of truth in music? It is not a truth related to facts of the world, nor is it a truth of the structure of music also. While there could be laws of auditory sensation, like the Gestalt laws of vision, these cannot be the truths of the aesthetic appreciation of music. In a trivial sense, one could accept truths of experience that arises from listening and responding to music. But we might hesitate to call these as truths since musical experience is first person and is immersed in the problem of subjective truths. All experience is truthful in this sense – that is 'really' happens to the individual and it may even be 'really true' that the individual experiences what she feels as her experience. But the aesthetic experience is more than experiences related to our 'ordinary' sensations.

Consider emotions. When I feel joy or sadness, it is indeed true that I am experiencing those emotions. These are psychological truths related to these emotions. We can describe these emotions; there are physiological and psychological consequences of these emotions. They influence our actions and our worldview. Thus, even in a pragmatic sense, these are truths. I am thinking of emotions here since music is very often seen to be associated with emotions. Music generates deep emotions in us. Indian classical music is particularly based on an intrinsic relation with emotion. The fact that different raagas are correlated with different emotions exemplifies this relation. Moreover, it is true that for many listeners (even when they are not classically trained or literate in these musical traditions) listening to a sad raaga evokes sadness or at least a particular kind of mood. And repeated experience of this relation consolidates this belief that certain musical expressions are conducive to certain kinds of experiences. Even non-classical music, ranging from pop to film songs, generate specific kinds of emotions.

One reason why we would resist discovering the idea of musical truth within these 'experiential truths' is that we have similar emotions which are catalysed by non-musical experiences. For example, we might feel sad on hearing some news or seeing an event in front of our eyes. In these cases, we believe that these events cause particular emotions. So one way to distinguish events which cause emotions and music which cause similar emotions is by recognizing that in the first case there is a causal relation whereas in the latter there is no causal relation. That is, music does not *cause* the emotions it does in the way events do. Neurobiologists might argue that the sounds of music might trigger certain chemicals which in turn will trigger the mechanisms leading to particular emotions. But the particular sounds or notes of music do not constitute music. They are the reduced components of a particular piece of music. The idea of music does not exist in these notes and their combination. Instead, it exists in a peculiar phenomenon, perhaps most evocatively described by Heidegger, that which is called 'presencing'. Music arises as a particular form of presence. Music emerges from these notes and their combination but it is not reducible to these elements. (The simplest way of recognizing this is to string together notes – very often this leads to the very opposite of music!) While Heidegger considers any artwork as 'the becoming and happening of truth', it is particularly so in the case of music. And there is also a difference in the happening or presencing of truth in music compared to other artworks, say, visual art.

First of all this truth which is presenced in music is not propositional. It does not communicate truths about any other aspect such as the truth of the performance, its relation to emotion, factual truth related to the world and so on. Levinson, for example, suggests that there are different aspects of truth that can be found in music such as structural correspondence between a piece of music and emotion it might correspond to or express. His conclusion relates truth to emotions, even though there may be different layers in this relation.

However, the truths of music are far more complex than their relation to emotions. Music cannot be reduced to an instrumental cause for emotions, even though that may often be the case. I began with the example of emotions primarily to find an entry into reflecting on the idea of truth that is inherent in every individual experience. There has been enough literature, spanning centuries, which have argued that such individual experiences are subjective and cannot stand the test of truth, which, by definition, has to be accessible to others. Even the debates on self-knowledge and individual intuition and insight have had to constantly respond to this problem. Does the case of music transcend these issues?

First of all, music as a form of presencing is independent of the emotions it creates. The emotions associated with music are second or higher order experiences of music. One can give a causal narrative of how certain pieces of music trigger certain memories which then trigger certain emotions and so on, and thus it is not the music per se that triggers these emotions.

The basic point is that – independent of the emotions triggered by music – we recognize something as music which is not merely a series of sounds. What is this capacity to recognize a new presence of music more than the collection of sounds? Is it merely a capacity of the human consciousness to integrate these discrete sounds into a unified whole? Or is there something in the very nature of certain combinations of sounds that create this sense of music?

The truth of music lies in this coming-to-be of that we call music. Emerging into presence is the fundamental character of music. So any notion of logic in music has to be discovered in this act of coming into being. The fundamental prerequisite of coming into being is the idea of truth. What comes into being is not a set of sounds or even a unified 'musical sound' for these are all material manifestations. What can come into being is only truth which makes us recognize music as music.

What is so special to truth in music? I want to suggest the following: Truth in music does not have a correlate of false – there is nothing false in music. Notes may be offkey or a combination of notes may sound jarring. But these are not about falsity of music. In fact, music gives us a formulation of truth with no notion of not-true. It is a logic of truth that is based on one value alone. Truth in art, in general, is a monovalent term; it is the only term that has no possibility of bivalence. Another way of expressing this is to note that in general all terms can either be true or false in bivalent logic. But what about 'truth' itself – can it take the values of true or false? (This might seem a classic example of category mistake but in the context of truth in music there is only the idea of truth which is co-present with the recognition of music.)

In the context of the 'logic' of science and literature which I mentioned in the beginning of this essay, we would have to phrase the logic of music as 'True and Never False'. There is an immediate problem with this claim. Suppose a particular musical phrase which, for example, characterizes particular raaga in Hindusthani music, is modified or even mistakenly sung. Would we then say that there is something wrong, something false about that? To this question, I would say that mistakes in the order of notes in a phrase or loss of shruti are mistakes in performance and are not really about anything being false in music. It is just plainly not music. In fact, such a performance does not let music come into presence and thus does not let the truth of music be made available to the listener. Thus, while we can even accept binaries such as music and not-music, it is not possible, as far as music is concerned, to have truth and not-truth.

Why do I want to invoke this notion of truth in music? Primarily it is to explore the aesthetics of logic. And why aesthetics in logic? That is very important because the very idea of formal logic, based as it is on forms, is about aesthetics. When we discover structure in logical arguments it could very often be an aesthetic judgement. Look at the examples from science – new theoretical discoveries, discovered through an aesthetic judgement are then converted to the theoretical. Recognition of new structures through the visual form of mathematical symbols is often an aesthetic judgement and this is part of the logical and not contrary to the logical.

Moreover, I want to draw attention to the importance of experience and feeling in logic, mathematics as well as in music. 'Feeling right' is an important notion in logic, mathematics and music. This sense of 'feeling right' often makes scientists choose particular structures. While they might validate it later, and in so doing reduce their initial insight as an 'aesthetic' judgement, this sense of 'feeling right' is so important to science and music. It is true that we do not necessarily have to invoke the idea of truth in the context of music but there is something profoundly truthful about music. But through this detour into music, what I want to place before Mihirda is this: logic has a world of 'feeling' associated with it. One way to understand it is to discover the aesthetic within the logical.

Sundar Sarukkai
Manipal Centre for Philosophy and Humanities,
Manipal University, India.
sarukkai1@yahoo.com

Contractifiability and Equivalent Metrices

Amita Sen

1 Introduction

The notion contractifiability is based on equivalent metrices. An operation $*$ is created on non-negative real members under which both the notions are preserved. In this paper we deal with the notion of contractifiability for which we require equivalent metric. The concept contractifiability is defined as follows –

Definition 1. *A continuous self-mapping T is said to be contractifiable in a metric space (X, d) if and only if there exists an equivalent metric d_1 such that T is contraction in (X, d_1).*

The foremost author regarding contractifiability is C. Bessaga (1957). His statement is "Let S be any arbitrary Set and $T : S \mapsto S$ be a mapping with the property that each iteration of T has a unique fixed point. Then, for each $\alpha \in (0, 1)$, there exists a metric d on S such that (S, d) is complete and T satisfies the inequality

$$d(Tx, Ty) \leq \alpha d(x, y) \text{ for all } x, y \in S.$$

Afterwards, in 1967 L. Janos established his Theorem :

Theorem 1. *If (X, d) is a compact metric space and $T :: X \mapsto X$ be a continuous mapping with the property*

$$\cap_0^\alpha T^n(X) = \{x_0\},\ x_0 \in X$$

then for each $\alpha \in (0, 1)$ there exists a metric D such that (X, D) is a compact space and T is a contraction mapping with respect to D.

P.R. Meyers (1984) has established the following equivalence –

i) T is contractifiable in (X, d)
ii) there exists $w \in X$ such that
 a) $Tw = w$
 b) $T^n x \mapsto w$, $x \in X$
 c) $T^n U = \{w\}$ for some neighbourhood U of w.

S. Leader (1977) has offered more simplified conditions which are –

a) $T^n x \mapsto w$, $x \in X$
b) $T^n x \mapsto w$ uniformly for all x in some neighbourhood U of x.

In this background we understand the requirement of finding a metric d^1 equivalent to a given metric d in the metric space (X, d) such that $T : X \mapsto X$ is a contraction mapping in (X, d_1). We will now introduce an operation $*$ on non-negative real numbers such that two metrices d_1 and d_2 on the same set, when connected by $*$ becomes a metric. Some useful properties of $*$ are obtained. Also, equivalence and contractifiability are preserved under the operation $*$.

2 Contractifiability and a binary operation

We define $*$ as follows:

Definition 2. *$*$ is a binary operation on non-negative real numbers satisfying the following axioms :*
i) $u * 0 = k_1 u$, $k_1 > 0$
ii) $0 * v = k_2 v$, $k_2 > 0$
and iii) $u \leq u_1 + u_2$, $v \leq v_1 + v_2$
$\Rightarrow u * v \leq (u_1 * v_1) + (u_2 * v_2)$

Corollary 1. $0 * 0 = 0$

Ex 1. $u * v = c_1 u + c_2 v$, $c_i > 0$, $i = 1, 2$
Ex 2. $u * v = \max\,(u, v)$
Ex 3. $u * v = (c_1 u^p + c_2 v^p)^{\frac{1}{p}}$, $c_i > 0$, $i = 1, 2$

Some of the interesting properties are:

P_1: $(u_1 + u_2) * v \leq u_1 * v + u_2 * v$ [Right Semi distributive]
P_2: $u * (v_1 + v_2) \leq u * v_1 + u * v_2$ [Left Semi distributive]
P_3: $u \leq u_1$ and $v \leq v_1 \Rightarrow u * v \leq u_1 * v_1$ [Monotone]
P_4: $k_1 u, k_2 v \leq u * v \leq k_1 u + k_2 v$

Also it can be proved that,

Theorem 2. *The topology on a set X induced by the metric $d_1 * d_2$ is finer than the topology induced by the metric d_1, or d_2.*

Definition 3. *Let d_1 and d_2 be two metrics defined on the sane set X. We define a function d by*
$d(x, y) = d_1 * d_2(x, y)$
$= d_1(x, y) * d_2(x, y)$

Theorem 3. *d is a metric on X but the converse is not true.*

Definition 4. *Let (X, d_1) and (X, d_2) be two metric spaces. If the identity mapping $I : (X, d_1) \mapsto (X, d_2)$ be a homeomorphism then d_1 and d_2 are said to be equivalent on X. We will write $d_1 \simeq d_2$.*

We now prove the following important theorem.

Theorem 4. *If $d_1 \simeq d_2$, then $d_1, d_2 \simeq d_1 * d_2$.*

Proof. Let $\{x_n\}$ be any sequence which converges to x with respect to d_1 i.e. $lim_{n\to\propto} d_1(x_n, x) = 0$
Since $d_1 \simeq d_2$, $lim_{n\to\propto} d_2(x_n, x) = 0$
We can show $lim_{n\to\propto} d_1 * d_2(x_n, x) = 0$ by P_4
Conversely let a sequence $\{y_n\}$ converge to y with respect to $d_1 * d_2$,
i.e. $lim_{n\to\propto} d_1 * d_2(y_n, y) = 0$
Then it can be shown
$lim_{n\to\propto} d_1(y_n, y) = 0$ [by P_4]
Hence $d_1 \simeq d_1 * d_2$. Similarly, $d_2 \simeq d_1 * d_2$.

Corollary 2. *Putting* $d_1 = d_2 = d$ *we can say* $d \simeq d * d$

Definition 5. *A word of metrics* d_1, d_2 *on the same set* X *is defined recursively by*

d_i, $i = 1, 2$ is a word
if d and d' are words then $(d * d')$ is a word.

Theorem 5. *If* $d_1 \simeq d_2$*, then any two words are equivalent to each other.*

Definition 6. *Two metrics d_1 and d_2 on the same set X are said to be uniformly equivalent if and only if the identity mapping $I : (X, d_1) \mapsto (X, d_2)$ and the inverse mapping $I^{-1} : (X, d_2) \mapsto (X, d_1)$ are uniformly continuous.*

Theorem 6. *If* $d_1 \simeq d_2$*, then each is* $\simeq d_1 * d_2$.

Proof. We take any arbitrary $\varepsilon > 0$ and $0 < \delta_1 < \varepsilon$ so that $\varepsilon - \delta_1 > 0$.
Since $d_1 \simeq d_2$, $I : (X, d_1) \mapsto (X, d_2)$ is uniformly continuous.
For $\frac{\delta_1}{k_2}$ there exists δ_2 such that for all $x, y \in X$,
$d_1(x, y) < \delta_2 \Rightarrow d_2(x, y) < \frac{\delta_1}{k_2}$
$\Rightarrow k_2 d_2(x, y) < \delta_1$.
Let $\delta = \min\,[\delta_2, \frac{\varepsilon - \delta_1}{k_1}]$
Then $d_1(x, y) < \delta \Rightarrow k_1 d_1(x, y) + k_2 d_2(x, y)$
$< k_1 . \frac{\varepsilon - \delta_1}{k_1} + \delta_1 = \varepsilon$
Hence by P_4, $d_1 * d_2(x, y) < \varepsilon$ i.e. $I : (X, d_1) \mapsto (X, d_1 * d_2)$ is uniformly continuous.
Again given $\varepsilon > 0$ we take $0 < \delta < k_1\varepsilon$, $d_1 * d_2(x, y) < \delta \Rightarrow d_1 * d_2(x, y) < k_1\varepsilon$
$\Rightarrow k_1 d_1(x, y) < k_1\varepsilon \Rightarrow d_1(x, y) < \varepsilon$.
So, $I^{-1} : (X, d_1 * d_2) \mapsto (X, d_1)$ is uniformly continuous.
Hence $d_1 \simeq d_1 * d_2$.
As uniform equivalence is transitive, $d_2 \simeq d_1 * d_2$.

Corollary 3. $d_1 = d_2 = d$ *implies* $d \simeq d * d$ *for any metric* d.

Repeating this process we can have a number of metrices which are equivalent to each other for a given metric d. We will define the set of all such equivalent metrices as $E_q(d)$. Now we will prove the relationship between $*$ and contractifiability.

Theorem 7. *Let a self mapping T be contractifiable in both the metric spaces (X, d_1) and (X, d_2). Then T is so in $(X, d_1 * d_2)$.*

Proof is obvious.

Corollary 4. *Under the given conditions T is contractifiable in (X, D) where D is any word of d_1, d_2 [Definition 5].*

Corollary 5. *Let $d_1 = d_2 = d$. Then we get, if T be contractifiable in (X, d), then T is so in $(X, d * d)$ and in (X, D) where D is any word of d.*

Corollary 6. *If T is contractifiable only in (X, d_2) and if metric d_1 does not exceed d_2, then also T is so in $(X, d_1 * d_2)$.*

Also we can establish the fact that —

Theorem 8. *Let T be contractifiable in $(X, d_1 * d_2)$ and let $d_2 \leq d_1$ then T is contractifiable in (X, d_1).*

The above theorems and the corollaries prove the importance of the operation $*$. We again return to the notion of contractifiability. The search for an equivalent metric d‘ in which the self mapping T is a contraction mapping is also easily available with the help of $*$. We now establish this fact.

Theorem 9. *Let i) T be a contraction mapping in a metric space (X, d) with contraction constant α ii) $\alpha < \frac{k_i}{k_1+k_2}, i = 1, 2$*

Then T is a contraction mapping in $(X, d * d)$.

Proof. Given that $d(Tx, Ty) \leq \alpha d(x, y)$ for all $x, y \in X,\ 0 < \alpha < 1$.

Then $d * d(Tx, Ty) \leq (k_1 + k_2)d(Tx, Ty)$ [by P_4] $\leq (k_1 + k_2)\alpha d(x, y) \leq \frac{k_1+k_2}{k_i}\alpha d * d(x, y)$ by P_4 , $i = 1$ or2.

This implies that T is a contraction mapping in $(X, d_1 * d_2)$.

Theorem 10. *Let i) T be a contraction mapping in (X, d_1) and*
ii) $d_i \leq d_1,\ i = 1, 2, ...(n + 1)$

Then T is a contraction mapping in X with all possible metrics formed by $d_i,\ i = 1, 2...(n + 1)$ and by different $*j,\ j = 1, 2.....n$ with same restrictions of k's and α.

The above results confirm the existence of a number of metrices which are all equivalent and the self mapping T remains a contraction mapping in all such new metric spaces. We return to the procedure of S. Leader who has found an equivalent metric in the following way.

Assumptions: Let $T : X \mapsto X$ be a mapping in (X, d).
i) T is continuous,
ii) $T^n x \to p$ with respect to the metric d and
iii) $T^n x \to p$ uniformly in some neighbourhood B of p with respect to d.

First he has taken $d'(x, y) = \min(1, d(x, y))$
Secondly he has constructed $\overline{d}(x, y) = \text{Sup}_{k>0}(d'(T^k x, T^k y))$
(This $\overline{d}$ becomes contractive)
Next he has taken sets $B_n = T^n B$ for all integrs n.
Finally he has constructed the largest pseudo metric D such that $D \leq \alpha^n \overline{d}$ on B_n i.e. $D(x, y) \leq \alpha^n\ \text{Sup}_{k>0}(d'(T^k x, T^k y))$

D is then a metric such that $D(Tx, Ty) \leq \alpha D(x, y)$ in (X, D).

It can be shown [3] that there are metrices D which are $\simeq d$ and a mapping T is contraction (X, D) but D cannot be found by Leader's procedure. To verify it we present a counter example.

Example. Let $D(x,y) = |x-y|$. X is the set of reals. $T : x \to \frac{x}{3}$ and $d(x,y) = \min(1, D(x,y))$.
Then $d \simeq D$ and $d \leq D$.
Also T is a contraction mapping in (X, D).
Since $d(T_{10}, T_1) = 1 = d(10,1)$, T is not contraction in (X, d).

We shall now prove that D can not be obtained from Leader's procedure.

$d'(x,y) = \min(1, d(x,y)) = min(1, min(1, D(x,y))) = min(1, D(x,y)) = d(x,y)$
$\overline{d}(x,y) = \text{Sup}_{k\geq 0}(d'(T^k x, T^k y)) = \text{Sup}_{k\geq 0}(d(T^k x, T^k y))$.
Now if D is the largest pseudo metric satisfying
$D \leq \alpha^n \overline{d}$ on B, for some set B we shall get
$D(x,y) \leq \alpha^n \text{Sup}_{k\geq 0}(d(T^k x, T^k y))$ on B_n,
i.e., $D(x,y) < \text{Sup}_{k\geq 0}(d(T^k x, T^k y))$ on B_n.
But for all $x, y \in X$, $\text{Sup}_{k\geq 0}(d(T^k x, T^k y)) = \text{Sup}(d(x,y), d(Tx,Ty), d(T^2x, T^2y), ...)$
$\leq \text{Sup}(D(x,y), D(Tx,Ty), ...) = D(x,y)$ as T is contraction in (X, D).

Hence D can not be found by Leader's method. This example shows that we can avoid Leader's procedure to find an equivalent metric in which T is contraction whereas with the help of the operation $*$ we can obtain a word of metrices which are all equivalent and T is contraction with respect to all such metrices.

References

1. Bessaga C.: On the converse of Banach's fixed point principle. Colloq Math 7 pp 41-43 (1959).
2. Janos L.: A converse of Banach's contraction principle. Proc. Amer, Math Soc., 18 (1967).
3. Sen A.: Some Topics on Metric Spaces related with Contraction Mapping. Ph.D thesis. University of Calcutta, (1993).
4. Lahiri B.K., Chakraborty M.K., Sen A. : Converse of Banach's contraction principle and star operation. Proc. Nat. Acad. Sci. India Sect. A, Vol 79, Pt IV (2009).
5. Leader S. : A topological characterisation of Banach's contraction. Pacific J. Mats. 69 pp 461-466 (1977)
6. Meyers P.R. : A converse of Banach's contraction theorems. J. Res. Bur. Standards 71B pp 73-76 (1967)
7. Meyers P.R. : On contractifiable self-mappings in non-linear analysis. Worls Scientific Pub. Co., Singapore, pp 407-432 (1987)

Amita Sen
Retired Faculty,
M.U.C Women's college, West Bengal, India.

A glimpse of linear logic and its algebra

Jayanta Sen

To enter into the world of linear logic, the paper "Linear logic" of J.-Y. Girard [9] and the book "Lecture on linear logic" by A. S. Troelstra [18] are very helpful. One may start with the two chapters of the book — on sequent calculus presentation and on algebraic semantics of linear logic.[1]

Starting from the materials presented in the above, we can now begin an investigation on sequent calculus presentation of linear logic and its algebraic structure.

The algebra intuitionistic linear logic is IL-algebra and that for classical linear logic is CL-algebra.

Definition. An intuitionistic linear algebra (IL-algebra) is a structure

$< X, \wedge, \vee, \perp, \multimap, *, 1 >$ where X is a non-empty set and $\wedge, \vee, \multimap, *$ are binary operations and $\perp, 1$ are two designated elements of X satisfying for all $x, y, z, x', y' \in X$ the conditions
(i)$< X, \wedge, \vee, \perp >$ is a lattice with least element $\perp$
(ii)$< X, *, 1 >$ is a commutative monoid with unit 1
(iii)$x \leq y$ and $x' \leq y'$ imply $x * y \leq x' * y'$ and $x' \multimap y \leq x \multimap y'$
(iv)$x * y \leq z$ iff $x \leq y \multimap z$, where $\leq$ is the lattice ordering.

It follows that the element $\perp \multimap \perp$ is the greatest element and is denoted by $\top$. It may be noted that for any $a \in X$, $1 \leq a \multimap a \leq \top$.

An IL-algebra X with a specified constant 0 is called classical linear algebra (*CL-algebra*) if $(x \multimap 0) \multimap 0 = x$ for all $x \in X$. Here, $\sim$ and $+$ can be defined as $\sim x := x \multimap 0$ and $x + y :=\sim (\sim x* \sim y)$. In a Cl-algebra, it is easy to check that $x \multimap y =\sim (x* \sim y)$.

It may be observed that in a CL-algebra $< X, \wedge, \vee, \perp, \multimap, *, 0, 1 >$ the following properties hold, for all $a, b, c \in X$. [4, 13]

1. $a * (b \vee c) = (a * b) \vee (a * c)$
2. $a \multimap (b \multimap c) = (a * b) \multimap c$
3. $\perp \multimap a = \top$
4. $a * b \leq (a * \top) \wedge (b * \top)$
5. If $1 \leq a, b$ then $a \vee b \leq a * b$
6. If $a, b \leq 1$ then $a * b \leq a \wedge b$
7. $a \leq a * \top$, equality holds if and only if $a \multimap a = \top$
8. $a * (a \multimap b) \leq b$
9. $(a \multimap b) * (b \multimap c) \leq a \multimap c$
10. $a * \perp = \perp$
11. $a \multimap \top = \top$
12. If $1 \leq a * \top$ then $a * \top = \top$
13. $(a * \top) * \top = a * \top$
14. If $a \leq b \leq a * \top$ then $a * \top = b * \top$
15. If $\top \neq 1$ then $a * \top \neq 1$

16. If $1 \leq a$ then $a \multimap \perp = \perp$ and $a * \top = \top$
17. $(a \multimap a) \multimap \perp = \perp$
18. $(a \multimap a) * \top = \top$
19. If $a \leq b$ then $a \multimap (b * \top) = \top$
20. $\sim a * a = \sim b * b$ does not hold, in general
21. $\sim \top = \perp$
22. $a \wedge b = \sim (\sim a \vee \sim b)$
23. $a \vee b = \sim (\sim a \wedge \sim b)$
24. $a \multimap b = \sim b \multimap \sim a$
25. $\sim 0 = 1$
26. $< X, +, 0 >$ is a commutative monoid with unit 0
27. $a + b = \sim a \multimap b$
28. $a + (b \wedge c) = (a + b) \wedge (a + c)$
29. $a + \perp = a$ if and only if $a * \top = a$
30. $a + \perp \leq a \leq a * \top$
31. $a + \perp = a * \top$ if and only if $a \multimap a = \top$
32. $a * \top \leq 0$ if and only if $a = \perp$
33. $\{a \in X : a * \top = a\} = \{a \in X : a + \perp = a\}$

The above properties of CL-algebra helps to find connection between CL-algebra and very similar MV-algebra.

An MV-algebra (many-valued algebra) [3, 5–7, 10] is a system $\mathbf{A} \equiv (A, +, *, \sim, 0, 1)$ obeying the following axioms, for all $x, y \in A$

1. $(A, *, 1)$ is a commutative monoid with unit 1
2. $x * 0 = 0$
3. $\sim\sim x = x$
4. $\sim 0 = 1$
5. $x * \sim x = 0$
6. $x + y = \sim (\sim x * \sim y)$
7. $x \vee y = y \vee x$, where $x \vee y := (x * \sim y) + y$.

It can be proved that an MV-algebra is a bounded lattice with least 0, greatest 1 and the ordering is defined by $x \leq y$ if and only if $x \vee y = y$.

A CL-algebra and an MV-algebra have many similarities. Both are bounded residuated lattices. But whereas in the first, the lattice is independent of the monoid only being linked through some axioms, in the second, the ordering is totally defined by the monoidal product and the negation. Although both are bounded, MV-algebras are integral residuated lattices, but CL-algebras are not so.

All the MV-algebras embedded in a CL-algebra can be characterized in the following sense.

A subset of a CL-algebra is an **embedded MV-algebra** if and only if it is an MV-algebra relative to the same operations $*, \sim$ (and hence $+$) and has the same ordering as the CL-algebra.

For a CL-algebra $\mathbf{X} \equiv (X, \wedge, \vee, \perp, \multimap, *, 1, 0)$, $(X, +, *, \sim, \perp, \top)$ is an MV-algebra if and only if for all $a, b \in X$, $a \cup b = b \cup a$, where $\sim a = a \multimap 0$ and $a \cup b = (a * \sim b) + b$.

Characterization all the MV-algebras embedded in a CL-algebra:—

For $A \subseteq X$ and $\theta, I \in X$, $\mathbf{M} \equiv (A, +, *, \sim, I)$, is an MV-algebra embedded in the CL-algebra $(X, \wedge, \vee, \bot, \multimap, *, 1, 0)$ if and only if

1. $\sim \theta = I$
2. $\theta \leq I$, where $' \leq '$ is the lattice-ordering of the CL-algebra
3. $A \subseteq \{x : \theta \leq x \leq I\}$
4. A is closed relative to $*$ and $\sim$
5. $x* \sim x = \theta$, for all $x \in A$ and
6. $x \cup y = y \cup x$, for all $x, y \in A$, where $x \cup y = (x* \sim y) + y$.

To find maximal embedded MV-algebra in a CL-algebra, if it exists, let us define **stem of a CL-algebra** $S(X)$ as $S(X) = \{x \in X : x * \top = x\}$.

$(S(X), +, *, \sim, \bot, \top)$ is an MV-algebra if and only if $x \cup y = y \cup x$, for all $x, y \in S(X)$.

Here it may be mentioned that stem is not always an MV-algebra.[2]

If $S(X)$ is an MV-algebra, it is a maximal one in the sense that it includes as subalgebras all MV-algebras with $\bot$ as the least element.

$(A, +, *, \sim, 0, 1)$ is an MV-algebra embedded in a CL-algebra $(X, \wedge, \vee, \bot, \multimap, *, 1, 0)$ if and only if

1. $0 \leq 1$
2. $A \subseteq \{x : 0 \leq x \leq 1\}$
3. A is closed relative to $*, \sim$ and
4. $x \cup y = y \cup x$, for all $x, y \in A$.

In addition, let $0 * 0 = 0, B = \{x : 0 \leq x \leq 1\}$. Let also that $x \cup y = y \cup x$, for all $x, y \in B$ then $(B, +, *, \sim, 0, 1)$ is an MV-algebra embedded in $(X, \wedge, \vee, \bot, \multimap, *, 1, 0)$.

Then a natural question is about the effects of these algebraic connections in corresponding logics. For that, a sequent calculus for multiplicative additive linear logic(MALL) (corresponding to CL-algebra) and a sequent calculus for Łukasiewicz many-valued logic($Ł_{\aleph_0}$) (corresponding to MV-algebra) are defined.

MALL: The language consists of the alphabet : variables p_i, constants $\top, 1$ and logical symbols $\sim, *$ and $\cup$. A, B stand for formulae. Sequents are of the form $\Gamma \Rightarrow \Delta$, where Γ, Δ are finite multisets (possibly empty) of formulae. $A + B :=\sim (\sim A* \sim B), \bot :=\sim \top, 0 :=\sim 1$ and $A \multimap B :=\sim (A* \sim B)$.

$Ax \qquad A \Rightarrow A$

$Cut \qquad \dfrac{\Gamma \Rightarrow A, \Delta \qquad \Gamma', A \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}$

$L\sim \qquad \dfrac{\Gamma \Rightarrow A, \Delta}{\Gamma, \sim A \Rightarrow \Delta} \qquad\qquad R\sim \qquad \dfrac{\Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \sim A, \Delta}$

$L* \qquad \dfrac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A*B \Delta} \qquad\qquad R* \qquad \dfrac{\Gamma \Rightarrow A, \Delta \qquad \Gamma' \Rightarrow B, \Delta'}{\Gamma, \Gamma' \Rightarrow A*B, \Delta, \Delta'}$

$L\cup \qquad \dfrac{\Gamma, A \Rightarrow \Delta \qquad \Gamma, B \Rightarrow \Delta}{\Gamma \Rightarrow A \cup B, \Delta} \qquad\qquad R\cup \qquad \dfrac{\Gamma \Rightarrow A, \Delta}{\Gamma \Rightarrow A \cup B, \Delta} \qquad \dfrac{\Gamma \Rightarrow B, \Delta}{\Gamma \Rightarrow A \cup B, \Delta}$

$L1 \quad \frac{\Gamma \Rightarrow \Delta}{\Gamma, 1 \Rightarrow \Delta}$ $\qquad R1 \quad \Rightarrow 1$

$RT \quad \Gamma \Rightarrow \top, \Delta$

$\text{Ł}_{\aleph_0}$: [15] The language consists of the alphabet : variables p_i and logical symbols $\sim$ and $*$. A, B, C etc. are formulae. Sequents are of the form $\Gamma \Rightarrow \Delta$, where Γ, Δ are finite multisets (possibly empty) of formulae. $A + B :=\sim (\sim A* \sim B)$ and $A \rightarrow B :=\sim (A* \sim B)$.

The axioms and rules are now stated.

$Ax \quad A \Rightarrow A$

$Cut \quad \frac{\Gamma \Rightarrow A, \Delta \qquad \Gamma', A \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}$

$LW \quad \frac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}$

$L\sim \quad \frac{\Gamma \Rightarrow A, \Delta}{\Gamma, \sim A \Rightarrow \Delta}$ $\qquad R\sim \quad \frac{\Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \sim A, \Delta}$

$L* \quad \frac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A*B \Delta}$ $\qquad R* \quad \frac{\Gamma \Rightarrow A, \Delta \qquad \Gamma' \Rightarrow B, \Delta'}{\Gamma, \Gamma' \Rightarrow A*B, \Delta, \Delta'}$

$LUB \quad \frac{A \Rightarrow C \qquad B \Rightarrow C}{(A \rightarrow B) \rightarrow B \Rightarrow C}$

Theorem. In MALL, if we add the following rule :

$LUB \quad \frac{A \Rightarrow C \qquad B \Rightarrow C}{(A \multimap B) \multimap B \Rightarrow C}$

then $\text{Ł}_{\aleph_0}$ is valid in the new system, say MALL$^+$ (i.e., MALL with LUB).

But for the embedding in logic, another kind of characterization of the algebraic embedding might be helpful. Boolean algebras are embedded in CL-algebra and the results are obtained for corresponding logics too. These are obtained by introducing a modal operator. In the same fashion, by a ξ-operator, MV-algebras can be embedded in a CL-algebra. [13, 15]

Let $\mathbf{X} \equiv (X, \wedge, \vee, \perp, \multimap, *, 1, 0)$ be a CL-algebra.
Then $\mathbf{Y} \equiv (\xi(X), +, *, \sim, \xi(\perp), \xi(\top))$ is an embedded MV-algebra, where ξ is a self-mapping of X and satisfies the following.

1. $\xi(X)$ is closed relative to $*, \sim$
2. $\xi(\perp) \leq \xi(x) \leq \xi(\top)$
3. $1 \leq \xi(\top)$
4. $\xi(x) * \xi(y) \leq \xi(x)$
5. $\xi(x) \leq \xi(z)$ and $\xi(y) \leq \xi(z)$ imply $(\xi(x) \multimap \xi(y)) \multimap \xi(y) \leq \xi(z)$.

Let us take ξ as the identity map. Then from the above result it follows that a CL-algebra becomes an MV-algebra if it satisfies the following conditions. $x * y \leq x$ and $x \leq z,\ y \leq z$ imply $(x \multimap y) \multimap y \leq z$. Observe that an alternative set has already been proved.

Let $\xi(x) = x * \top$ for all $x \in X$. Then $\xi(X)$ becomes the stem of X. Here also $\xi(X)$ is an MV-algebra if it satisfies some conditions — which are proved in the first discussion of embedding issues.

Converse of the result involving ξ-operator also holds good.

For every embedded MV-algebra $\mathbf{Y} \equiv (Y, +, *, \sim, \theta, I)$ in a CL-algebra $\mathbf{X} \equiv (X, \wedge, \vee, \perp, \multimap, *, 1, 0)$, there is $\xi : X \to X$ satisfying all the conditions of the previous result, more precisely, $\mathbf{Y} = \xi(X)$.

It may be noted that the mapping ξ is not unique. In fact, $\xi(y) = y$ for all $y \in Y, \xi(\perp) = \theta, \xi(\top) = I$ are necessary conditions and for all other $x \in X, \xi(x)$ may be any element of Y.

This ξ-operator helps to introduce a modal operator in the language of η in MALL and extend the logic suitably to get a new system, say , MALL_η, although η is not exact logical counterpart of ξ.

MALL_η is MALL with the following extra axiom and rules.

$Ax \qquad 1 \Rightarrow \eta(\top)$

$LW_\eta \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}$

$LUB_\eta \qquad \frac{A \Rightarrow C \qquad B \Rightarrow C}{(A \to B) \to B \Rightarrow C}$

for all formulae A, B, C for which η occurs exactly before all its atomic constituents.

The definition of the interpretation for the language of $\text{Ł}_{\aleph_0}$ into that of MALL_η is given by

$p^0 := \eta(p)$, for non-constant atomic p

$I^0 := \eta(\top)$

$(\sim A)^0 :=\sim A^0$

$(A * B)^0 := A^0 * B^0$.

Obviously, for any formula A of $\text{Ł}_{\aleph_0}$, in A^0, η occurs exactly before each non-constant atomic constituents and each occurrence of I in A is replaced by $\eta(\top)$.

Theorem. If $\vdash \Gamma \Rightarrow \Delta$ in $\text{Ł}_{\aleph_0}$ then $\vdash \Gamma^0 \Rightarrow \Delta^0$ in MALL_η.

For the converse part, the forgetful interpretation of MALL_η into $\text{Ł}_{\aleph_0}$ is defined as follows.

$v(\top) = v(I) = I$

$v(A * B) = v(A) * v(B)$

$v(A \cup B) = (v(A) \to v(B)) \to v(B)$

$v(\sim A) =\sim v(A)$

$v(\eta(A)) = v(A)$

$v(p) = p$, for non-constant atomic p.

So, $v(A \multimap B) = v(A) \to v(B)$.

Theorem. If $\vdash \Gamma \Rightarrow \Delta$ in MALL_η then $\vdash v(\Gamma) \Rightarrow v(\Delta)$ in $\text{Ł}_{\aleph_0}$.

Lemma. $\vdash v(\Gamma)^0 \Rightarrow v(\Delta)^0$ in $\text{Ł}_{\aleph_0}$ if and only if $\vdash \Gamma \Rightarrow \Delta$ in $\text{Ł}_{\aleph_0}$.

Theorem. $\vdash \Gamma \Rightarrow \Delta$ in $\text{Ł}_{\aleph_0}$ if and only if $\vdash \Gamma^0 \Rightarrow \Delta^0$ in MALL_η.

One very important, but different branch is rough set theory. tqBa [1, 2] is an important algebra in the study of rough sets.[3] Interestingly, a distributive CL-algebra(i.e., a CL-algebra with the extra condition that the lattice is distributive) can be transformed into a topological quasi-Boolean algebra(tqBa).

Definition. An algebra $< X, \leq, \wedge, \vee, \sim, L, \perp, \top >$ is a topological quasi-Boolean algebra(tqBa) if and only if for all $a, b \in A$ the following conditions hold.

1. $< X, \leq, \wedge, \vee, \sim, \perp, \top >$ is a quasi-Boolean algebra(qBa), i.e.,
   (a) $< X, \leq, \wedge, \vee, \sim, \perp, \top >$ is a distributive lattice with least $\perp$ and greatest $\top$
   (b) $\sim\sim a = a$
   (c) $\sim (a \vee b) = \sim a \wedge \sim b$
2. $La \leq a$
3. $L(a \wedge b) = La \wedge Lb$
4. $LLa = La$
5. $L\top = \top$
6. $\sim L \sim La = La$.

Theorem. If $< X, \wedge, \vee, \perp, \multimap, *, 0, 1 >$ is a distributive CL-algebra, then $< X, \leq, \wedge, \vee, \sim, L, \perp, \top >$ is a tqBa, where '$\leq$' is the lattice-ordering, $\sim a := a \multimap 0$, $La := a + \perp$ and $\top$ is the greatest element of the lattice.

For the logical counterpart of the result, a sequent calculus **tqBl** for tqBa is defined. [17]

The language consists of the alphabet : variables p_i, constants $\top, \perp$ and logical symbols $l, \sim, \cap$ and $\cup$. A, B etc. stand for formulae. Sequents are defined in the usual sense.

$Ax1 \quad A \Rightarrow \sim\sim A \qquad Ax2 \quad \sim\sim A \Rightarrow A \qquad Ax3 \quad \sim l \sim lA \Rightarrow lA$

$Cut \quad \dfrac{\Gamma \Rightarrow A, \Delta \qquad \Gamma', A \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}$

$Rule \sim \quad \dfrac{\Gamma \Rightarrow \Delta}{\sim\Delta \Rightarrow \sim\Gamma}$

$LW \quad \dfrac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta} \qquad RW \quad \dfrac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow A, \Delta}$

$LC \quad \dfrac{\Gamma, A, A \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta} \qquad RC \quad \dfrac{\Gamma \Rightarrow A, A, \Delta}{\Gamma \Rightarrow A, \Delta}$

$L\cup \quad \dfrac{\Gamma, A \Rightarrow \Delta \qquad \Gamma', B \Rightarrow \Delta'}{\Gamma, \Gamma', A \cup B \Rightarrow \Delta, \Delta'} \qquad R\cup \quad \dfrac{\Gamma \Rightarrow A, B \Delta}{\Gamma \Rightarrow A \cup B, \Delta}$

$L\cap \quad \dfrac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A \cap B \Rightarrow \Delta} \qquad R\cap \quad \dfrac{\Gamma \Rightarrow A, \Delta \qquad \Gamma' \Rightarrow B, \Delta'}{\Gamma, \Gamma' \Rightarrow A \cap B, \Delta, \Delta'}$

$L\perp \quad \Gamma, \perp \Rightarrow \Delta \qquad R\top \quad \Gamma \Rightarrow \top, \Delta$

$Ll \quad \dfrac{\Gamma, A \Rightarrow \Delta}{\Gamma, lA \Rightarrow \Delta} \qquad Rl \quad \dfrac{l\Gamma \Rightarrow A}{l\Gamma \Rightarrow lA}$

Let '^{0}' be the translation from the well-formed formulae of tqBl to those of dMALL(i.e., MALL with the additional axiom

$(dAx) \qquad A \cap (B \cup C) \Rightarrow (A \cap B) \cup (A \cap C)$)

given by

$p^0 := p$, for atomic p

$(\sim A)^0 := \sim A^0$

$(A \cup B)^0 := A^0 \cup B^0$

$(A \cap B)^0 := A^0 \cap B^0$

$(lA)^0 := A^0 + \perp$.

Hence, $(\sim l \sim A)^0 := A^0 * \top$.

Theorem. If $\vdash \Gamma \Rightarrow \Delta$ in tqBl then $\vdash (\Gamma^0) \Rightarrow [\Delta^0]$ in dMALL, where (Γ^0) means the lattice conjuncts of the formulae of Γ^0 and $[\Delta^0]$ means the lattice disjuncts of the formulae of Δ^0. Also, $(\Gamma^0) = \top$ if Γ is empty and $[\Delta^0] = \bot$ if Δ is empty.

The sequent calculus tqBl also helps to define sequent calculus for all the algebraic structures lattice, bounded lattice, bounded lattice with negation, quasi-Boolean algebra without distributivity, quasi-Boolean algebra and also for pre-rough algebra and topological Boolean algebra.

Let us go back to CL-algebra. There are four constants, viz., $\bot, 0, 1$ and $\top$. $\bot$ and $\top$ are the least and the greatest elements. But what are the positions of 0 and 1? Although three possibilities are there, viz., $0 \leq 1$, $1 \leq 0$ and 0,1 are incomparable, but $0 \leq 1$ is the most interesting case. 0 can be thought of the first false and after that one gradually approaches downwards to absolute false $\bot$. Parallely, 1 can be considered as the truth and then one gradually approaches upwards to absolute truth $\top$.

The algebraic result 'if $1 \leq a, b$ then $a * b \geq a \vee b$' — is very rare and it looks like that after getting minimal truth, interaction of two formulae goes above their join. This result may provoke one to further restrict the structure that the part $1 \leq x \leq \top$ is linear(equivalently, $\bot \leq x \leq 0$ is linear).

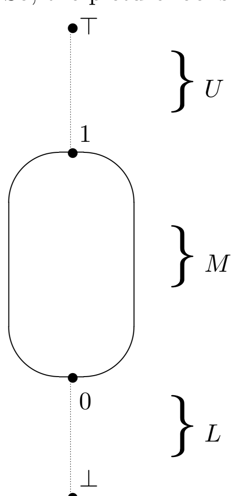
If the part $0 \leq x \leq 1$ is an MV-algebra , then it may be considered as an extension of MV-algebra. [14]

Definition. An extended many-valued algebra 1 (EMVA1) $\mathcal{X} \equiv < X, \wedge, \vee, \bot, \multimap, *, 1, 0 >$ is a CL-algebra where $0 \leq 1$ and $< M, *, \sim, 0, 1 >$ is an MV-algebra, where $M = \{x \in X : 0 \leq x \leq 1\}$.

Definition. An extended many-valued algebra 2 (EMVA2) is an EMVA1 with the additional property that the underlying set $X = L \cup M \cup U$, where $L = \{x : \bot \leq x \leq 0\}$ and $U = \{x : 1 \leq x \leq \top\}$ and L, U are linear.

The extended many valued algebras EMVA1 and EMVA2 are extensions of many valued algebras in the sense that these algebras reduce to MV-algebra if $\top = 1 (\bot = 0$, equivalently).

So, the picture looks like



For the sequent calculus EMVL1 corresponding to the algebra EMVA1, all the axioms and rules of MALL are taken. Further the following axiom and rule are added.

$A1 \quad 0 \Rightarrow 1$ $\qquad LUB \quad \frac{0\Rightarrow A \quad A\Rightarrow 1 \quad 0\Rightarrow B \quad B\Rightarrow 1}{(A\multimap B)\multimap B\Rightarrow (B\multimap A)\multimap A}$

MV-algebras are now considered as a suitable algebra for representing vagueness. The extended MV-algebras may prove to be even better algebras for the purpose.

---

[1] Dr. Mohua Banerjee had a good collection of some very important papers on linear logic. It helped me to study on it. Also, within few months of my Ph.D. work, Dr. Purander Bhaduri gave a series of lectures on linear logic at Calcutta Logic Circle conference. That too helped me a lot.

[2] My Ph. D. supervisor Prof. Mihir Kumar Chakraborty first noticed the connections between MV-algebra and CL-algebra. The name STEM is also given by him. According to Prof. Chakraborty, stem may play very important role in the study of CL-algebra. We got some very important results on stem, but did not go through it into a great length.

[3] The algebra tqBa was introduced by Prof. Chakraborty and Dr. Banerjee in their study on rough sets.

References

1. M. Banerjee and M.K. Chakraborty : *Rough algebra*, Bull Polish Acad Sc (Math), 41(4):293-297, 1993.
2. M. Banerjee and M.K. Chakraborty : *Rough sets through algebraic logic*, Fundamenta Informaticae 28(3-4): 211-221, 1996.
3. L.P. Belluce, A. di Nola and S. Sessa : *Triangular norms, MV-algebras and bold fuzzy set theory*, Math Japonica, 36:481-487, 1991.
4. M.K. Chakraborty and J. Sen : *MV-algebras embedded in a CL-algebra*, Int Jr Approximate Reasoning, 18 : 217-229, 1998.
5. C.C. Chang : *Algebraic analysis of many-valued logics*, Tran. Am. Math. Soc., 88, 467-490, 1958.
6. C.C. Chang : *A new proof of the completeness of the Łukasiewicz axioms*, Tran. Am. Math. Soc., 93, 74-80, 1959.
7. R.L.O. Cignoli, I.M.L. D'Ottaviano and D. Mundici : *Algebraic foundations of many-valued reasoning*, Kluwer Acad Pub, 2000.
8. N. Galatos, P. Jipsen, T. Kowalski and H. Ono : *Residuated lattices: an algebraic glimpse at substructural logics*, Elseiver, 2007.
9. J.-Y. Girard : *Linear logic*, Theoretical Computer Science, 50, 1-102, 1987.
10. D. Mundici : *Interpretation of AF C^*-algebras in Łukasiewicz sentential calculus*, Jr Funct Anal, 65, 15-63, 1986.
11. H. Ono : *Structural rules and a logical hierarchy*, in P.P. Petkov(ed.), Math Logic, Plenum Press, 95-104, 1990.

12. H. Ono : *Substructural logics and residuated lattices — an introduction*, in 50 yrs of Studia Logica, 21 of Trends in Logic: 193-228, Kluwer Acad Publ, 2003.
13. J. Sen : *Some embeddings in linear logic and related issues*, Ph.D. dissertation, University of Calcutta, India, 2000.
14. J. Sen and M.K. Chakraborty : *Logic on Extended MV-algebras*, Proc Int Conf Communications, Devices and Intelligent Systems, CODIS,Jadavpur University, Kolkata, India, 2004.
15. J. Sen and M.K. Chakraborty *Linear logic and Łukasiewicz* $\aleph_0$*-valued logic: a logico algebraic study*, Jr of applied non-classical logics, 11: 313-329, 2001.
16. J. Sen and M.K. Chakraborty *Fuzzy Logic on Affine CL-Algebra*, ed by D. Chakraborty, S. Nanda and D. Mazumder, in Fuzzy Logic and Its Application to Technology and Management, Narosa Pub House, 2007.
17. J. Sen and M.K. Chakraborty *A study of interconnections between rough and Łukasiewicz 3-valued logic*, Fundamenta Informaticae 51: 311 324, 2002.
18. A.S. Troelstra : *Lectures on linear logic*, No. 29, CSLI, Stanford, 1992.

Jayanta Sen
Department of Mathematics,
Presidency University, Kolkata, India.
jsen_01@yahoo.co.in

Some aspects of Fuzzy Hyperstructure Theory⋆

M.K. Sen

1 Introduction

Hyperstructure theory was born in 1934 when Marty [11] defined hypergroups, began to analysis their properties and applied them to groups, rational algebraic functions. Now they are widely studied from theoretical point of view and applied them to many subjects of pure and applied mathematics [28, 32]. In 1965 L.A.Zadeh [18] introduced the notion of a fuzzy subset of a non-empty set X, as a function from X to $[0, 1]$. In 1971 A.Rosenfeld [1] defined the concept of fuzzy group. Since then many papers have been published in the field of fuzzy algebra (see Mordeson et al.)[14].

Recently fuzzy set theory has been well developed in the context of hyperalgebraic structure theory. The study of fuzzy hyperstructures is an interesting research topic of fuzzy sets. There is a considerable amount of work on the connections between fuzzy sets and hyperstructures. This work can be divided into three main approaches. The first approach defines crisp hyperoperations through fuzzy sets. This was initiated by Corsini [28] and continued by himself [32] and others. The second approach concerns fuzzy hyperalgebras which is a direct generalization of fuzzy algebras (fuzzy (sub) groups, fuzzy lattices, fuzzy rings, etc.). This approach can be extended to fuzzy hypergroups. For example, given a crisp hypergroup $(H, \cdot)$ and a fuzzy subset μ of H, then we say that μ is a fuzzy (sub) hypergroup of $(H, \cdot)$ if every cut $\mu_t (t \in [0, 1])$ of μ is a (crisp) sub hypergroup of $(H, \cdot)$. This was initiated by Zahedi et al. [22] and further studied by Ameri et al. [33], Yamak et al. [36], Zhan et al. [16] and so on. The third approach involves something which is also called fuzzy hypergroup, but it is completely different from what we described above. It was introduced by Corsini and Tofan [30]. The basic idea is the following: a crisp hyperoperation assigns to every pair of elements a crisp set; a fuzzy hyper operation assigns to every pair of elements a fuzzy set. This idea was continued in Kehagias [2], Sen et al. [24, 25, 26], Leoreanu [38] and others. In 1967 Wee [39] introduced the concept of fuzzy automata. Since then fuzzy automata theory has been developed by many researchers. Malik, Mordeson and Sen introduced the concept of Fuzzy finite state machine and a fuzzy transformation semigroup based on Wee's concept of fuzzy automata and studied different properties in the papers [4 - 8].The theory of Hypercompositional is being introduced into the Theory of Fuzzy Automata and Fuzzy Languages by G. Massouros [3] and Sen et al.[25, 27].

Let H be a non-empty set. An equivalence relation on H is a binary relation on

⋆ On 65th birthday of Prof. M.K. Chakraborty

H that is reflexive, symmetric and transitive. Every equivalence relation gives rise to a partition on the underlying set H. Formally, a partition of a set H is a collection P of pairwise disjoint subsets of H whose union is equal to H. Again, every partition of H gives rise to an equivalence relation on H. Let us denote the set of all equivalence relations by $E(H)$ and the class of all partitions on H by $P(H)$. Then, we have the following relations:
Let $R \in E(H)$. For every $a \in H$, let $R_a = \{x \in H : (a, x) \in R\}$. Then $P_R = \{R_a : a \in H\}$ is a partition of H. Let P be a partition of H. Define a relation R_P by $R_P = \{(x, y) \in H \times H :$ there exists a subset Q of H such that $\{x, y\} \subseteq Q \in P\}$. Then $R_P \in E(H)$. It can be shown that there exists a one-to-one correspondence between $E(H)$ and $P(H)$ given by $R \to P_R$.

A fuzzification of the above concepts, viz., fuzzy equivalence relation and fuzzy partition, have been dealt with in many works. In fact, the first such definition was proposed by Zadeh himself [19], wherein he proposed the following definition, which is now termed as similarity relations.

Definition 1. *A mapping $\rho : H \times H \longrightarrow [0, 1]$ is said to be a fuzzy relation on H. ρ is said to be fuzzy reflexive if $\rho(x, x) = 1$ for all $x \in H$. ρ is said to be fuzzy symmetric if $\rho(x, y) = \rho(y, x)$ for all $x, y \in H$. ρ is said to be fuzzy transitive if $\rho \circ \rho \leq \rho$, where $(\rho \circ \rho)(x, y) = \vee\{\rho(x, z) \wedge \rho(z, y) : z \in H\}$ for all $x, y \in H$.*

A fuzzy relation ρ on a set H is called a fuzzy equivalence relation on H if it is fuzzy reflexive, fuzzy symmetric and fuzzy transitive.

Following Zadeh [19], many works have appeared generalising the above concept of a fuzzy equivalence relation. For instance, many types of reflexivity have been proposed and discussed, see for instance, [20, 21, 17]. Similarly, other types of transitivity have also been discussed.

The concept of a fuzzy partition is a generalisation of the parition of a set H. Ruspini [9] was the first to propose one such generalisation. To work dealing with either different or more general definitions see, for instance, [13 - 21, 23, 37]. The following is an equivalent form of the above definitions that is more or less established in the literature.

Definition 2. *A collection P of fuzzy subsets of a set X is called a fuzzy partition of X if the following properties are satisfied:*
(i) For all $U \in P$ there is some $x \in X$ such that $U(x) = 1$,
(ii) For all $x \in X$ there is exactly one $U \in P$ such that $U(x) = 1$,
(iii) If $U, V \in P$ such that $U(x) = V(y) = 1$ for some $x, y \in X$, then $U(y) = V(x)$.

Recently Jayaram Balusubramaniam and Mesiar Radko [15] have studied the fuzzyfication of the equivalence relation and also studied the one-to-one correspondence between the fuzzy equivalence relation and fuzzy partition.

In this article my attempt will be to concentrate mainly on some of our recent works on Fuzzy Hyper Structure Theory.

2 Fuzzy hyperaction and fuzzy hyperset

A (left) action of a monoid S on a nonempty set Q is a mapping $f : S \times Q \longrightarrow Q$ (usually denoted by $f\ (x, q) {\longrightarrow} xq$) for all $x \in S$ and $q \in Q$. The set Q is called an S-set if there exists an action of S on Q such that
(i) $(xy)q = x(yq)$ for all $x, y \in S$ and $q \in Q$ and (ii) $1q = q$ for all $q \in Q$.

This concept of S-set plays an important role in the theory of Deterministic finite automata [16]. Considering the theory of Non-deterministic finite automata [16] Sen and Chowdhury [37] have introduced the concept of S-hyperset as a generalization of S-set.

Throughout the section S denotes a monoid, Q denotes a nonempty set, $P(Q)$ denotes the set of all subsets of Q and $F(Q)$ denotes the set of all fuzzy subsets of Q.

Definition 3. *A (left) hyperaction of S on Q is a mapping $\circ : S \times Q \mapsto P(Q)$ (usually denoted by $\circ(x, q) \mapsto x \circ q$) for all $x \in S$ and $q \in Q$.*
Let $A \in P(Q)$ and $x \in S$. We define $x \circ A \in P(Q)$ by

$$x \circ A = \begin{cases} \cup\{x \circ a : a \in A\}, & \text{if } A \neq \emptyset \\ \emptyset, & \text{if } A = \emptyset \end{cases}$$

A set Q is called an S-hyperset of type 1 if there exists a hyperaction $\circ$ of S on Q such that (i) $(xy) \circ q = x \circ (y \circ q)$ for all $x, y \in S$ and $q \in Q$ and (ii) $q \in 1 \circ q$ for all $q \in Q$.
A set Q is called an S-hyperset of type 2 if there exists a hyperaction $\circ$ of S on Q such that (i) $(xy) \circ q = x \circ (y \circ q)$ for all $x, y \in S$ and $q \in Q$ and (ii) $1 \circ q = \{q\}$ for all $q \in Q$.

If a set Q is an S-hyperset of type 1 or of type 2 with respect to a hyperaction $\circ$ then we write the hyperset as $(Q, \circ)$.
In [25] we have defined fuzzy hyperaction of a monoid on a set and studied different properties.

Definition 4 (25). *A $(left)$ fuzzy hyperaction of S on Q is a function $\bullet$: $S \times Q \longrightarrow F(Q)$ (usually denoted by $\bullet(x, q) \longrightarrow x \bullet q$) for all $x \in S$ and $q \in Q$.*

Let $q \in Q$, $\mu \in F(Q)$ and $x \in S$. We now define $S \bullet q$, $x \bullet \mu$, $S \bullet \mu \in F(Q)$ as follows:
$(S \bullet q)(p) = \vee\{(x \bullet q)(p) : x \in S\}$, $\quad (x \bullet \mu)(p) = \vee\{(x \bullet r)(p) \wedge \mu(r) : r \in Q\}$
and

$$\begin{aligned}(S \bullet \mu)(p) &= \vee\{(S \bullet r)(p) \wedge \mu(r) : r \in Q\} = \vee\{(x \bullet r)(p) \wedge \mu(r) : x \in S \text{ and } r \in Q\} \\ &= \vee\{(x \bullet \mu)(p) : x \in S\}.\end{aligned}$$

Definition 5 (25). *A set Q is called an S-fuzzy hyperset of type 1 if there exists a fuzzy hyperaction $\bullet$ of S on Q such that*
(i) $x \bullet (y \bullet q) = xy \bullet q$ for all $x, y \in S$ and $q \in Q$ and (ii) $(1 \bullet q)(q) = 1$ for all $q \in Q$.

Definition 6 (25). *A set Q is called an S-fuzzy hyperset of type 2 if there exists a fuzzy hyperaction $\bullet$ of S on Q such that*
(i) $x \bullet (y \bullet q) = xy \bullet q$ for all $x, y \in S$ and $q \in Q$ and
(ii) $(1 \bullet q)(p) = \begin{cases} 1 \text{ if } p = q \\ 0 \text{ if } p \neq q \end{cases}$ for all $p, q \in Q$.

If a set Q is an S-fuzzy hyperset of type 1 or of type 2 with respect to a fuzzy hyperaction $\bullet$ then we write the fuzzy hyperset as $(Q, \bullet)$.

For any S-hyperset $(Q, \circ)$ there is associated an S-fuzzy hyperset $(Q, \bullet)$, called the associated S-fuzzy hyperset of the S-hyperset $(Q, \circ)$, where the fuzzy hyperaction $\bullet$ of S on Q is defined by $(x \bullet q)(p) = \begin{cases} 1, \text{ if } p \in x \circ q \\ 0, \text{ otherwise} \end{cases}$ *for all $x \in S$ and $q, p \in Q$.*

In [29] we have shown that there exists a correspondence between a $\Sigma^\star$-fuzzy hyperset and a fuzzy finite state machine $\overline{\mathcal{M}} = (\Sigma, Q, \mu)$.

Theorem 1. *Let Σ and Q be two nonempty finite sets. $\Sigma^\star$ denotes the free monoid generated by Σ. Let $\bullet$ be a fuzzy hyperaction of $\Sigma^\star$ on Q and satisfying the properties $x \bullet (a \bullet q) = xa \bullet q$ for all $x \in \Sigma^\star$, $a \in \Sigma$, $q \in Q$ and*
$(\lambda \bullet q)(p) = \begin{cases} 1, \text{ if } p = q \\ 0, \text{ otherwise} \end{cases}$ for all $q, p \in Q$, where λ is the identity element of $\Sigma^\star$ i.e. the empty word. Then $(Q, \bullet)$ is a $\Sigma^\star$-fuzzy hyperset.

Definition 7 (5). *A fuzzy finite state machine fuzzy finite state machine is a 3-tuple $\overline{\mathcal{M}} = (\Sigma, Q, \mu)$, where Σ is a nonempty finite set, called the set of alphabets, Q is a nonempty finite set, called the set of states, $\mu : Q \times \Sigma \times Q \longrightarrow [0, 1]$, called the fuzzy transition function of the fuzzy finite state machine $\overline{\mathcal{M}}$. The extended transition function $\mu^\star : Q \times \Sigma^\star \times Q \longrightarrow [0, 1]$ of μ is defined by*
$\mu^\star(p, \lambda, q) = \begin{cases} 1, \text{ if } p = q \\ 0, \text{ otherwise} \end{cases}$ for all $p, q \in Q$ and
$\mu^\star(p, xa, q) = \vee\{\mu^\star(p, x, r) \wedge \mu(r, a, q) : r \in Q\}$ for all $x \in \Sigma^\star$, $a \in \Sigma$ and $p, q \in Q$.

Lemma 1. *Let $\overline{\mathcal{M}} = (\Sigma, Q, \mu)$ be a fuzzy finite state machine and $\mu^\star$ be the extension of μ. Then $\mu^\star(p, xy, q) = \vee\{\mu^\star(p, x, r) \wedge \mu^\star(r, y, q) : r \in Q\}$ for all $x, y \in \Sigma^\star$ and $p, q \in Q$.*

Theorem 2. *Let $\overline{\mathcal{M}} = (\Sigma, Q, \mu)$ be a fuzzy finite state machine. Then $(Q, \bullet)$ is a $\Sigma^\star$-fuzzy hyperset, where the fuzzy hyperaction $\bullet$ of $\Sigma^\star$ on Q is given by $(x \bullet q)(p) = \mu^\star(p, x, q)$ for all $x \in \Sigma^\star$ and $q, p \in Q$.*

Theorem 3. *Let Σ and Q be two nonempty finite sets and $(Q, \bullet)$ be a $\Sigma^\star$-fuzzy hyperset. Then there is a fuzzy finite state machine $\overline{\mathcal{M}} = (\Sigma, Q, \mu)$ such that the extension $\mu^\star : Q \times \Sigma^\star \times Q \longrightarrow [0, 1]$ of μ is given by*
$\mu^\star(p, \lambda, q) = \begin{cases} 1, \text{ if } p = q \\ 0, \text{ otherwise} \end{cases}$ for all $p, q \in Q$ and
$\mu^\star(p, xa, q) = \vee\{\mu^\star(p, x, r) \wedge \mu(r, a, q) : r \in Q\}$ for all $x \in \Sigma^\star$, $a \in \Sigma$ and $p, q \in Q$.

3 Fuzzy hyperset and fuzzy regular language

In Section 2 we have defined hyperset and fuzzy hyperset. In this section we define fuzzy regular language in terms of fuzzy hyperset and study some properties of it which are already studied in terms of fuzzy finite automata.

Definition 8. *A fuzzy subset $\overline{L}$ of $\Sigma^\star$ is called a fuzzy regular language on Σ if there exists a $\Sigma^\star$-fuzzy hyperset $(Q, \bullet)$ and an element $q_0 \in Q$ and a fuzzy subset μ of Q such that $\overline{L}(x) = \vee\{(x \bullet q_0)(p) \wedge \mu(p) : p \in Q\}$ for all $x \in \Sigma^\star$.*

Definition 9. *A subset L of $\Sigma^\star$ is called a regular language on Σ if there exists a $\Sigma^\star$-hyperset $(Q, \circ)$, an element $q_0 \in Q$ and a nonempty subset F of Q such that $L = \{x \in \Sigma^\star : x \circ q_0 \cap F \neq \phi\}$.*

Theorem 4. *Let L be a regular language on Σ. Then there is a $\Sigma^\star$-hyperset $(Q, \circ)$ and an element $q_0 \in Q$ and a nonempty subset F of Q such that $L = \{x \in \Sigma^\star : x \circ q_0 \cap F \neq \phi\}$. Let $(Q, \bullet)$ be the associated $\Sigma^\star$-fuzzy hyperset of the $\Sigma^\star$-hyperset $(Q, \circ)$. Let χ_F be the characteristic function of F. Then the fuzzy subset χ_L, characteristic function of L can be written as $\chi_L(x) = \vee\{(x \bullet q_0)(p) \wedge \chi_F(p) : p \in Q\}$ for all $x \in \Sigma^\star$.*

From the above Theorem it follows that if L is a regular language on Σ then χ_L is a fuzzy regular language on Σ.

Theorem 5. *Let χ_L, the characteristic function of L, be a fuzzy regular language on Σ. Then there is a $\Sigma^\star$-fuzzy hyperset $(Q, \bullet)$ and an element $q_0 \in Q$ and a fuzzy subset μ of Q such that $\chi_L(x) = \vee\{(x \bullet q_0)(p) \wedge \mu(p) : p \in Q\}$ for all $x \in \Sigma^\star$. Define a mapping $\circ : \Sigma^\star \times Q \longrightarrow P(Q)$ by $x \circ q = \{p \in Q : (x \bullet q)(p) = 1\}$ for all $x \in \Sigma^\star$ and $q \in Q$. Then $(Q, \circ)$ is a $\Sigma^\star$-hyperset and L can be written as follows:*
$L = \{x \in \Sigma^\star : x \circ q_0 \cap \mu_1 \neq \phi\}$, *where* $\mu_1 = \{p \in Q : \mu(p) = 1\}$.

From the above Theorem it follows that if the characteristic function χ_L of L is a fuzzy regular language on Σ then L is a regular language on Σ.

4 Fuzzy decomposition of fuzzy hyperset

Keeping in mind the definition and properties of fuzzy partition we study fuzzy decomposition of an S-fuzzy hyperset into sub fuzzy hypersets.

Definition 10. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. A fuzzy subset μ of Q is said to be a sub fuzzy hyperset if $x \bullet \mu \leq \mu$ for all $x \in S$.*

Definition 11. *An S-fuzzy hyperset $(Q, \bullet)$ is said to be connected if*
$\vee\{(x \bullet q)(p) : x \in S\} = 1$ *for all $q, p \in Q$.*

Definition 12. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. A sub fuzzy hyperset μ of Q is said to be proper if $|Im(\mu)| > 1$.*

Definition 13. *An S-fuzzy hyperset $(Q, \bullet)$ is said to be simple if it has no proper sub fuzzy hyperset.*

Definition 14. *By a fuzzy decomposition of an S-fuzzy hyperset $(Q, \bullet)$ we mean a fuzzy partition $P = \{\mu_i : i \in I(index\ set)\}$ of Q into sub fuzzy hypersets μ_i of Q. Then we write $\chi_Q = \vee\{\mu_i : i \in I\}$, called a fuzzy decomposition of Q. If no such decomposition is possible with $|I| > 1$ then Q is said to be indecomposable.*

Theorem 6. *Every connected S-fuzzy hyperset $(Q, \bullet)$ is indecomposable.*

Theorem 7. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. Let P, the collection of sub fuzzy hypersets of Q, form a fuzzy decomposition of $(Q, \bullet)$ into sub fuzzy hypersets of $(Q, \bullet)$. Then there is a fuzzy equivalence relation ρ on Q such that $(x \bullet q)(p) \leq \rho(q, p)$ for all $x \in S$ and $q, p \in Q$ and each member of P is a fuzzy ρ-equivalence class.*

Theorem 8. *Let $(Q, \bullet)$ be an S-fuzzy hyperset and ρ be a fuzzy equivalence relation on Q such that $(x \bullet q)(p) \leq \rho(q, p)$ for all $x \in S$ and $q, p \in Q$. Then for each $q \in Q$ the fuzzy equivalence class ρ_q is a sub fuzzy hyperset of Q.*

Theorem 9. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. Let ρ be a fuzzy equivalence relation on Q such that $(x \bullet q)(p) \leq \rho(q, p)$ for all $x \in S$ and $q, p \in Q$. Then the collection $P = \{\rho_p : p \in Q\}$ of fuzzy equivalence classes ρ_p forms a fuzzy decomposition of $(Q, \bullet)$ into sub fuzzy hypersets ρ_p of $(Q, \bullet)$.*

Theorem 10. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. We define a fuzzy relation τ on Q by $\tau(q, p) = (S \bullet q)(p)$ for all $q, p \in Q$. Then the fuzzy relation $\tau^\star = \vee\{(\tau \vee \tau^{-1})^n : n \geq 1\}$ is a fuzzy equivalence relation on Q such that $\tau \leq \tau^\star$ and satisfies the condition $(x \bullet q)(p) \leq \tau^\star(q, p)$ for all $q, p \in Q$ and $x \in S$. Also each $\tau^\star_p$ is a sub fuzzy hyperset of $(Q, \bullet)$.*

Theorem 11. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. Then the collection $P = \{\tau^\star_p : p \in Q\}$ forms a fuzzy decomposition of $(Q, \bullet)$ into sub fuzzy hypersets $\tau^\star_p$ of $(Q, \bullet)$.*

Definition 15. *A sub fuzzy hyperset μ of an S-fuzzy hyperset $(Q, \bullet)$ is said to be sub fuzzy co-hyperset of $(Q, \bullet)$ if $\mu(p) \wedge (x \bullet q)(p) \leq \mu(q)$ for all $q, p \in Q$ and $x \in S$.*

Theorem 12. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. Let $\tau^\star$ be the fuzzy equivalence relation on Q defined in the Theorem 5.10. Then for each $p \in Q$, the fuzzy equivalence class $\tau^\star_p$ is a sub fuzzy co-hyperset of $(Q, \bullet)$.*

Theorem 13. *Let μ be a sub fuzzy co-hyperset of an S-fuzzy hyperset $(Q, \bullet)$. Then $\mu(q) \wedge \tau^\star(q, p) \leq \mu(p)$ for all $q, p \in Q$.*

Theorem 14. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. Then for each $p \in Q$ the fuzzy equivalence class $\tau^\star_p$ is the smallest sub fuzzy co-hyperset of $(Q, \bullet)$ such that $\tau^\star_p(p) = 1$.*

Theorem 15. *Let $(Q, \bullet)$ be an S-fuzzy hyperset. Then the collection $P = \{\tau_p^\star : p \in Q\}$ forms a fuzzy decomposition of $(Q, \bullet)$ into smallest sub fuzzy co-hypersets $\tau_p^\star$ of $(Q, \bullet)$ such that $\tau_p^\star(p) = 1$ for all $p \in Q$.*

5 Fuzzy semihypergroup and fuzzy hypergroup

Definition 16. *In section 2 we have seen fuzzy hyperaction $\circ : S \times Q \longrightarrow F(Q)$ of a monoid S on a set Q. In [28] Sen, Ameri and Chowdhury have considered the fuzzy hyperaction $\bullet : H \times H \longrightarrow F(H)$ of a nonempty set H on H as fuzzy hyperoperation on H, where $F(H)$ is the set of all fuzzy subsets of H. Thus a fuzzy hyperoperation on H is a mapping $\bullet : H \times H \longrightarrow F(H)$ written as $(a, b) \longrightarrow a \bullet b$ for all $a, b \in H$.*
H together with a fuzzy hyperoperation $\bullet$ is called a fuzzy hypergroupoid and is denoted by $(H, \bullet)$.
Let $x, p \in H$ and $\mu, \nu \in F(H)$. Then we define $x \bullet \mu$, $\mu \bullet x$, $\mu \bullet \nu \in F(H)$ by $(x \bullet \mu)(p) = \vee\{(x \bullet r)(p) \wedge \mu(r) : r \in H\}$, $(\mu \bullet x)(p) = \vee\{(r \bullet x)(p) \wedge \mu(r) : r \in H\}$ and
$(\mu \bullet \nu)(p) = \vee\{\mu(r) \wedge (r \bullet s)(p) \wedge \nu(s) : r, s \in H\}$ for all $x, p \in H$.

Definition 17. *A fuzzy hypergroupoid $(H, \bullet)$ is called a fuzzy semihypergroup if $(a \bullet b) \bullet c = a \bullet (b \bullet c)$ for all $a, b, c \in H$.*

Definition 18. *A fuzzy semihypergroup $(H, \bullet)$ is called a fuzzy hypergroup if $a \bullet H = H \bullet a = \chi_H$ for all $a \in H$, where $a \bullet H$ and $H \bullet a$ are defined by $(a \bullet H)(p) = \vee\{(a \bullet r)(p) : r \in H\}$, $(H \bullet a)(p) = \vee\{(r \bullet a)(p) : r \in H\}$ for all $p \in H$ and*
$(a \bullet \chi_H)(p) = \vee\{(a \bullet r)(p) \wedge \chi_H(r) : r \in H\} = \vee\{(a \bullet r)(p) : r \in H\} = (a \bullet H)(p)$,
$(\chi_H \bullet a)(p) = \vee\{\chi_H(r) \wedge (r \bullet a)(p) : r \in H\} = \vee\{(r \bullet a)(p) : r \in H\} = (H \bullet a)(p)$ for all $p \in H$.

Some examples of fuzzy semihypergroups and fuzzy hypergroups.

Example 1. Let H be a nonempty set. Define a fuzzy hyperoperation $\bullet$ on H by $a \bullet b = \chi_{\{a,b\}}$ for all $a, b \in H$, where $\chi_{\{a,b\}}$ denotes the characteristic function of the set $\{a, b\}$. Let $a, b, c \in H$. Then
$((a \bullet b) \bullet c)(t) = \vee\{(a \bullet b)(r) \wedge (r \bullet c)(t) : r \in H\} = \vee\{\chi_{\{a,b\}}(r) \wedge \chi_{\{r,c\}}(t) : r \in H\}$
$= \begin{cases} 1, \text{ if } t \in \{a, b, c\} \\ 0, \text{ otherwise} \end{cases}$ for all $t \in H$.

Similarly we can show that $(a \bullet (b \bullet c))(t) = \begin{cases} 1, \text{ if } t \in \{a, b, c\} \\ 0, \text{ otherwise} \end{cases}$ for all $t \in H$.

Therefore $((a \bullet b) \bullet c)(t) = (a \bullet (b \bullet c))(t)$ for all $a, b.c, t \in H$. This shows that $(H, \bullet)$ is a fuzzy semihypergroup.
Now $(H \bullet a)(t) = \vee\{(x \bullet a)(t) : x \in H\} = \vee\{\chi_{\{x,a\}}(t) : x \in H\} = 1 = \chi_H(t)$ for all $t \in H$. Therefore $H \bullet a = \chi_H$ for all $a \in H$. Similarly we can show that $a \bullet H = \chi_H$ for all $a \in H$. Therefore $(H, \bullet)$ is a fuzzy hypergroup.

Example 2. Let S be a semigroup and $\mu(\neq 0)$ be a fuzzy subsemigroup of S. Let $a, b \in S$. Define a fuzzy hyperoperation $\bullet$ on S by
$(a \bullet b)(t) = \begin{cases} \mu(a) \wedge \mu(b), & \text{if t = ab} \\ 0, & \text{otherwise.} \end{cases}$
It can be shown that $(S, \bullet)$ is a fuzzy semihypergroup.

Example 3. Let μ be a fuzzy subgroup of an abelian group G. We define a fuzzy hyperoperation $\bullet$ on G by $(a \bullet b)(t) = \mu(abt^{-1})$. Then we can show that $(G, \bullet)$ is a fuzzy hypergroup.

In the following Proposition we show that from each S-fuzzy hyperset we can construct a fuzzy hypergroup.

Proposition 1. *Let $(Q, \bullet')$ be an S-fuzzy hyperset. Define a fuzzy hyperoperation $\bullet$ on Q by $(p \bullet q)(r) = (S \bullet' p)(r) \vee (S \bullet' q)(r)$ for all $p, q, r \in Q$. Then $(Q, \bullet)$ is a fuzzy hypergroup.*

Lemma 2. *Let $(H, \bullet)$ be a fuzzy semihypergroup. Let $a, b \in H$ and μ be a fuzzy subset of H. Then (i) $\chi_a \bullet \chi_b = a \bullet b$, (ii) $\chi_H \bullet \chi_a = H \bullet a$, (iii) $\chi_H \bullet \mu = H \bullet \mu$ and (iv) $\mu \bullet \chi_H = \mu \bullet H$.*

Using the Lemma we prove the following Theorem.

Theorem 16. *Let $(H, \bullet)$ be a fuzzy hypergroup. Then $a \bullet b \neq 0$ for all $a, b \in H$.*

6 Fuzzy semihypergroup associated with fuzzy relation

Sen and Chowdhury have introduced [26] fuzzy hyper groupoid H_ρ corresponding to a fuzzy relation ρ on H and studied H_ρ as fuzzy semihypergroup and fuzzy hypergroup imposing necessary and sufficient conditions on ρ.

Definition 19. *Let ρ be a fuzzy relation on a set H. Define a fuzzy hyperoperation $\bullet$ on H by $(a \bullet b)(x) = \rho(a, x) \vee \rho(b, x)$ for all $a, b, x \in H$. Then $(H, \bullet)$ is a fuzzy hypergroupoid and is denoted by $H_\rho = (H, \bullet)$. By definition of $\bullet$ we have $(a \bullet a)(x) = \rho(a, x)$ for all $a, x \in H$. Then $(a \bullet b)(x) = (a \bullet a)(x) \vee (b \bullet b)(x)$ for all $a, b, x \in H$. Also $(a \bullet b)(x) = (b \bullet a)(x)$ for all $a, b, x \in H$. Therefore $(H, \bullet)$ is a fuzzy commutative hypergroupoid.*

Recall that a fuzzy hypergroupoid $(H, \bullet)$ is said to be a fuzzy semihypergroup if $a \bullet (b \bullet c) = (a \bullet b) \bullet c$ for all $a, b, c \in H$, where the fuzzy subsets $a \bullet \mu$ and $\mu \bullet a$ of H are defined as follows: $(a \bullet \mu)(x) = \vee\{(a \bullet r)(x) \wedge \mu(r) : r \in H\}$ and $(\mu \bullet a)(x) = \vee\{\mu(r) \wedge (r \bullet a)(x) : r \in H\}$ for all $a, x \in H$. The following example of a fuzzy relation ρ on a set H shows that $H_\rho = (H, \bullet)$ is a fuzzy hypergroupoid but not a fuzzy semihypergroup.

Example 4. Let $H = \{1, 2, 3, 4, ...\}$, the set of all positive integers. Define a fuzzy relation ρ on H by $\rho(x, y) = \frac{1}{xy}$ for all $x, y \in H$. Then $H_\rho = (H, \bullet)$ is a fuzzy hypergroupoid. Consider the members $2, 4, 5, 3$ of H. Then $((2 \bullet 4) \bullet 5)(3) = \frac{1}{3}$ and $(2 \bullet (4 \bullet 5)(3) = \frac{1}{4}$. Thus $((2 \bullet 4) \bullet 5)(3) \neq (2 \bullet (4 \bullet 5))(3)$ and so H_ρ is not a fuzzy semihypergroup.

Lemma 3. *Let $a, b, c, x \in H_\rho$. Then*
(i) $((a \bullet b) \bullet c)(x) = (\rho \circ \rho)(a, x) \vee (\rho \circ \rho)(b, x) \vee (\vee\{\rho(a, r) \vee \rho(b, r) : r \in H\} \wedge \rho(c, x))$.
(ii) $(a \bullet (b \bullet c))(x) = (\rho(a, x) \wedge (\vee\{\rho(b, r) \vee \rho(c, r) : r \in H\})) \vee (\rho \circ \rho)(b, x) \vee (\rho \circ \rho)(c, x)$.

Theorem 17. *Let ρ be a fuzzy relation on a set H. Then following statements are equivalent:*
(i) $H_\rho = (H, \bullet)$ *is a fuzzy semihypergroup.*
(ii) $(\rho \circ \rho)(a, x) \leq (\rho \circ \rho)(b, x) \vee (\vee\{\rho(b, r) : r \in H\})$, $(\rho \circ \rho)(a, x) \leq \rho(a, x) \vee (\rho \circ \rho)(b, x)$ *and* $\rho(a, x) \leq (\rho \circ \rho)(a, x) \vee \rho(b, x)$ *for all* $a, b, x \in H$.
(iii) $(\rho \circ \rho)(a, x) \leq (\rho \circ \rho)(b, x) \vee (\vee\{\rho(b, r) : r \in H\})$ *and* $(\rho \circ \rho)(a, x) \vee \rho(b, x) = \rho(a, x) \vee (\rho \circ \rho)(b, x)$ *for all* $a, b, x \in H$.

If ρ is a fuzzy reflexive relation on H then $\rho(r, r) = 1$ for all $r \in H$. Consequently $\vee\{\rho(b, r) : r \in H\} = 1$ for all $b \in H$ and $(\rho \circ \rho)(a, x) \geq \rho(a, x) \wedge \rho(x, x) = \rho(a, x)$ for all $a, x \in H$. Then we have the following inequalities: $(\rho \circ \rho)(a, x) \leq (\rho \circ \rho)(b, x) \vee (\vee\{\rho(b, r) : r \in H\})$ and $\rho(a, x) \leq (\rho \circ \rho)(a, x) \vee \rho(b, x)$ for all $a, b, x \in H$. Thus we have the following Corollary:

Corollary 1. *Let ρ be a fuzzy reflexive relation on a set H. Then following statements are equivalent:*
(i) $H_\rho = (H, \bullet)$ *is a fuzzy semihypergroup.*
(ii) $(\rho \circ \rho)(a, x) \leq \rho(a, x) \vee (\rho \circ \rho)(b, x)$ *for all* $a, b, x \in H$.
(iii) $(\rho \circ \rho)(a, x) \vee \rho(b, x) = \rho(a, x) \vee (\rho \circ \rho)(b, x)$ *for all* $a, b, x \in H$.

Theorem 18. *Let ρ be a fuzzy reflexive and transitive relation on H. Then H_ρ is a fuzzy semihypergroup.*

Theorem 19. *Let ρ be a fuzzy relation on a set H such that H_ρ is a fuzzy semihypergroup. Then $\rho \circ \rho$ is fuzzy transitive.*

In the conclusion of this section we point out that the fuzzy hypergroupoid H_ρ introduced in the Definition 6.1 can be studied by defining the fuzzy subsets $a \bullet_\rho \mu$, $\mu \bullet_\rho a$ and $\mu \bullet_\rho \nu$ of H in H_ρ as follows: $(a \bullet_\rho \mu)(x) = \rho(a, x) \vee (\vee\{\mu(p) \wedge \rho(p, x) : p \in H\})$, $(\mu \bullet_\rho a)(x) = \vee\{\mu(p) \wedge \rho(p, x) : p \in H\} \vee \rho(a, x)$ and $(\mu \bullet_\rho \nu)(x) = \vee\{(\mu(p) \vee \nu(p)) \wedge \rho(p, x) : p \in H\}$ for all $a, x \in H$. Obviously $(\mu \bullet_\rho \mu)(x) = \vee\{\mu(p) \wedge \rho(p, x) : p \in H\}$. Then $(\mu \bullet_\rho \nu)(x) = (\mu \bullet_\rho \mu)(x) \vee (\nu \bullet_\rho \nu)(x)$ for all $x \in H$. Now $(\chi_a \bullet_\rho \mu)(x) = (\chi_a \bullet_\rho \chi_a)(x) \vee (\mu \bullet_\rho \mu)(x) = \vee\{\chi_a(p) \wedge \rho(p, x) : p \in H\} \vee (\mu \bullet_\rho \mu)(x) = \rho(a, x) \vee (\vee\{\mu(p) \wedge \rho(p, x) : p \in H\}) = (a \bullet_\rho \mu)(x)$ for all $x \in H$. Similarly we can show that $(\mu \bullet_\rho \chi_a)(x) = (\mu \bullet_\rho a)(x)$ for all $x \in H$.

Recall that a fuzzy semihypergroup $(H, \bullet)$ is said to be a fuzzy hypergroup if $(a \bullet H)(x) = (H \bullet a)(x) = 1$ for all $a, x \in H$, where the fuzzy subsets $a \bullet H$ and $H \bullet a$ of H are defined by $(a \bullet H)(x) = \vee\{(a \bullet r)(x) : r \in H\}$ and $(H \bullet a)(x) = \vee\{(r \bullet a)(x) : r \in H\}$ for all $a, x \in H$.

Theorem 20. *Let ρ be a fuzzy relation on a set H. Then H_ρ is a fuzzy hypergroup if and only if* *(i)* $(\rho \circ \rho)(a, x) \leq (\rho \circ \rho)(b, x) \vee (\vee\{\rho(b, r) : r \in H\})$, $(\rho \circ \rho)(a, x) \vee \rho(b, x) = \rho(a, x) \vee (\rho \circ \rho)(b, x)$ *for all* $a, b, x \in H$ *and (ii)* $\vee\{\rho(r, x) : r \in H\} = 1$ *for all* $x \in H$.

Theorem 21. *Let* H_ρ *be a fuzzy hypergroup. Then*
(i) ρ *is fuzzy symmetric* $\Longrightarrow \rho \circ \rho$ *is a fuzzy equivalence relation on* H *and*
(ii) ρ *is fuzzy symmetric and* $|H/(\rho \circ \rho)| > 1 \Longrightarrow \rho$ *is a fuzzy reflexive relation on* H.

Theorem 22. *Let* ρ *be a fuzzy reflexive and fuzzy transitive relation on* H*. Then* H_ρ *is a fuzzy hypergroup.*

Theorem 23. *Let* ρ_1 *and* ρ_2 *be two fuzzy relations on* H *such that both are fuzzy reflexive and transitive. Then* $H_{\rho_1 \wedge \rho_2}$ *is a fuzzy hypergroup.*

Definition 20. *Let* H_1 *and* H_2 *be two nonempty sets and* ρ_1, ρ_2 *be fuzzy relations on* H_1, H_2 *respectively. Define a fuzzy relation* ρ *on* $H_1 \times H_2$ *by* $\rho((a_1, a_2), (x_1, x_2)) = \rho_1(a_1, x_1) \wedge \rho_2(a_2, x_2)$ *for all* $a_1, x_1 \in H_1$ *and* $a_2, x_2 \in H_2$.

Theorem 24. *Let* ρ_1 *and* ρ_2 *be fuzzy relations on* H_1 *and* H_2 *respectively. Let* ρ *be the fuzzy relation on* $H = H_1 \times H_2$ *defined in the Definition 20 such that* H_ρ *is a fuzzy hypergroup. Then* $H_{1\rho_1}$ *and* $H_{2\rho_2}$ *are fuzzy hypergroups and* $\rho_1(a_1, x_1) \vee (\rho_2 \circ \rho_2)(b_2, x_2) = (\rho_1 \circ \rho_1)(a_1, x_1) \vee \rho_2(b_2, x_2)$ *for all* $a_1, x_1 \in H_1$ *and* $b_2, x_2 \in H_2$.

Theorem 25. *Let* ρ_1 *and* ρ_2 *be fuzzy reflexive relations on* H_1 *and* H_2 *respectively such that* $H_{1\rho_1}$ *and* $H_{2\rho_2}$ *are fuzzy hypergroups and* $\rho_1(a_1, x_1) \vee (\rho_2 \circ \rho_2)(b_2, x_2) = (\rho_1 \circ \rho_1)(a_1, x_1) \vee \rho_2(b_2, x_2)$ *for all* $a_1, x_1 \in H_1$ *and* $b_2, x_2 \in H_2$*. Then corresponding to the fuzzy relation* ρ *on* $H = H_1 \times H_2$ *defined in the Definition 20,* H_ρ *is a fuzzy hypergroup.*

Theorem 26. *Let* ρ *be a fuzzy reflexive relation on a set* H *such that* H_ρ *is a fuzzy hypergroup. Then* H_{ρ^n} *is a fuzzy hypergroup for all* $n = 1, 2, ...$

Theorem 27. *Let* ρ *be a fuzzy reflexive relation on a set* H *and* $|H| = n$*. Then* $(H, \bullet_{n-1})$ *is a fuzzy hypergroup.*

Theorem 28. *Let* $(H, \bullet)$ *be a fuzzy semihypergroup. Then there is a fuzzy relation* ρ *on* H *such that* $(H, \bullet)$ *is of the form* H_ρ *if and only if the following conditions are satisfied:*
(i) $a \bullet b = a^2 \vee b^2$, *(ii)* $a^2 \leq (a^2)^2$ *and (iii)* $(a^2)^2 \vee b^2 = a^2 \vee (b^2)^2$, *for all* $a, b \in H$, *where* a^2 *means the fuzzy subset* $a \bullet a$ *of* H *and* $(a^2)^2$ *means the fuzzy subset* $(a \bullet a) \bullet (a \bullet a)$ *of* H.

References

1. A.Rosenfeld, Fuzzy groups, J. Math. Anal. Appl. 35 (1971), 512–517.
2. A. Kehagias, K. Serafimidis, The L-fuzzy Nakano hypergroup, Inform. Sci. 169(2005)305–327.
3. C.G. Massouros, Hypercompositional Structures in the theory of languages and automata. Analele Stintifice Ale Universitatii AL. I.Cuza IASI Tomul III (1994), 65–73.
4. D. S. Malik, J.N.Mordeson, M.K.Sen, On subsystems of Fuzzy finite state machine. Fuzzy Sets and Systems. 68(1994), 83–92.
5. D. S. Malik, J.N.Moderson, M.K.Sen, Semigroups of Fuzzy infinite state machines. P.P.Wang, ed. Advances in Fuzzy Theory and Technology, Vol.II (1995), 87–98.
6. D. S. Malik, J.N.Mordeson, M.K.Sen, The Cartesian composition of Fuzzy machines. Kybernets. International Journal of Systems and Cybernetrics. 24 (1995), 167–179.
7. D. S. Malik, J.N.Mordeson, M.K.Sen, On Fuzzy regular languages. Information Sciences. 88 (1996), 263–273.
8. D. S. Malik, J.N.Mordeson, M.K.Sen, Admissible partitions of Fuzzy finite state machines. International Jour. of Uncertaintity, Fuzziness and Knowledge Based Systems. 5 (1997), 723–732.
9. E.H.Ruspini, A new approach to nclustering.Inform. Contr.15(1969) 22–32.
10. E.P. Klement, B. Moser, On the redundancy of fuzzy partitions, Fuzzy Sets Syst. 85 (1997) 195-201.
11. F.Marty, Sur une generalisation de la notion de groupe, Huiteme congres des math.scand.,Stockholm 1934,45-49.
12. I.G.Rosenberg, Hypergroups and Join spaces determined by relations, Italian Journal of Pure and Applied Mathematics, Vol., No.4, (1998) 93–101.
13. J.C. Bezdek, J.D. Harris, Fuzzy partitions and relations: an axiomatic basis for clustering, Fuzzy Sets Syst. 1 (1978) 111-127.
14. J.N.Mordeson, D.S.Malik, Fuzzy Commutative Algebra, World Scientific Publishing Co.Pte.Ltd,1998n
15. Jayaram Balasubramaniam, Mesiar Radka I-Fuzzy equivalence relation and I-fuzzy partition. Information Sciences 179(2009) 1278–1297
16. J. Zhan, B. Davvaz, K.P. Shum, On probabilistic n-ary hypergroups, Inform. Sci. 180 (2010) 1159–1166.
17. K.C. Gupta, R.K. Gupta, Fuzzy equivalence relations redefined, Fuzzy Sets Syst. 79 (1996) 227-233.
18. L.A.Zadeh, Fuzzy Sets, Inform. and Control, vol. 8 (1965), 338–353.
19. L.A.Zadeh, Similarity relations and fuzzy ordering. Inform.Sci.3 (1971) 177–200.
20. M.K. Chakraborty, S. Das, On fuzzy equivalence I, Fuzzy Sets Syst. 11 (1983) 185-193.
21. M.K.Chakraborty, S. Das, On fuzzy equivalence II, Fuzzy Sets Syst. 12 (1983) 299-307.
22. M.M.Zahedi, M. Bolurian, A. Hasankhani, On polygroups and Fuzzy subpolygroups, Journal of Fuzzy Mathematics, Vol., No.1, (1995) 1–15.

23. M. Cric, J. Ignjatovic, S. Bogdanovic, Fuzzy equivalence relations and their equivalence classes, Fuzzy Sets Syst. 158 (2007) 1295-1313.
24. M.K.Sen, Reza Ameri and G. Chowdhury, Fuzzy hypersemigroups.Soft Computing 12(9) (2008), 891–900.
25. M.K.Sen, G. Chowdhury and Reza Ameri, Fuzzy Hyperaction of Monoids and Fuzzy Finite State Machines. Bulletin of Calcutta Mathematical Society, 100(4) (2008), 359–378.
26. M.K.Sen, G. Chowdhury, Fuzzy Semihypergroup associated with Fuzzy relation. The Journal of Fuzzy Mathematics 18(1) 2010, 193–208.
27. M.K.Sen, G. Chowdhury, Regular and Fundamental relation on an S-hyperset, International Journal of Mathematical Sciences Vol. 10, No. 1-2, January-June 2011, pp. 79–89.
28. P.Corsini, Prolegomena of Hypergroup Theory, Aviani Editore (1993).
29. P. Corsini, V. Leoreanu, Join spaces associated with fuzzy sets, J. Comb., Inf. Syst. Sci. 20(1995)293–303.
30. P. Corsini, I. Tofan, On Fuzzy Hypergroups, PU.M.A., Vol. 8, (1997) 29–37.
31. P. Corsini, V. Leoreanu, Fuzzy sets and join spaces associated with rough sets, Rend. Circ. Mat. Palermo 51(2002)527–536.
32. P.Corsini, V. Leoreanu, Applications of Hyperstructure Theory, Kluwer Academic Publications (2003).
33. R.Ameri, M. M. Zahedi, Hypergroup and join spaces induced by a fuzzy subset, PU.M.A,Vol.8 (1997) 155–168.
34. S.K. Bhakat, P. Das, q-Similitudes and q-fuzzy partitions, Fuzzy Sets Syst. 51 (1992) 195-202.
35. S. Bodjanova, Complement of fuzzy k-partitions, Fuzzy Sets Syst. 62 (1994) 175-184.
36. S. Yamak, O. Kazanc, B. Davvaz, Applications of interval valued fuzzy n-ary polygroups with respect to t-norms (t-conorms), Inform. Sci. 178 (2008) 3957–3972.
37. V.Murali, Fuzzy equivalence relations, Fuzzy Sets Syst. 30(1989),155–161.
38. V. Leoreanu, Direct limit and inverse limit of join spaces associated with fuzzy sets, Pure Math. Appl. 11(2000)509–516.
39. W.G.Wee, On generalizations of adaptive algorithm and application of the fuzzy sets concept to pattern classification. Ph.D. Thesis. Purde Univ.,1967.
40. W.M.L. Holcombe, Algebraic automata theory, Cambridge University Press, 1982.

M.K. Sen
Retired Faculty,
Department of Pure Mathematics,
University of Calcutta, India.
senmk6@yahoo.com

The Nature of *ANUMEYA*: Some Early Indian Views[1]

Prabal Kumar Sen

I

Anumana (*i.e*. inference) has been acknowledged as an accredited source of knowledge or means of knowledge (***pramana***) by many schools of Indian philosophy, though there is considerable difference of opinion among these schools (and in some cases, even among the adherents of the same school) regarding the different issues regarding ***anumana*** (*eg*. the definition or nature of ***anumana***, the types of ***anumana***, the basis of ***anumana*** and so on). In inference (***anumana***), the object to be inferred (***anumeya*** / ***sadhya*** / ***gamya*** / ***lingi***) is established in a certain locus (***paksa***) on the basis of the fact that some inferential mark (***anupaksa*** / ***sadhana*** / ***hetu*** / ***gamaka*** / ***linga***); which is (i) pervaded by the ***sadhya*,** as well as (ii) present in the ***paksa***. Pervasion (***vyapti***) occurs between a ***hetu*** (say H) and a ***sadhya*** (say S) if it so happens that *either* or *both* of the following conditions obtain:

i) Wherever H is present, S is also present,

ii) Wherever S is absent, H is also absent.

If pervasion (***vyapti***) obtains between H and S, then H is said to be *pervaded* (***vyapya***), and S is said to be its pervader (***vyapaka***). The stock example of ***anumana***, where the presence of ***sadhya*** is established in the ***paksa*** on the basis of the presence of ***hetu*** in that ***paksa*** is where one infers the existence of fire in a distant hill from the fact that an unbroken column of smoke is arising from that hill.

It should be mentioned here that according to the Indian thinkers who admit ***anumana*** as a means of knowledge, there are certain situations in which an ***anumana*** that is supposed to establish the existence of an entity in a certain locus does not, or cannot take place. The first one is where the object or entity that is supposed to be inferred in a certain locus has already been known with certainty (***siddha***) in that very locus through some other means of knowledge. Thus if we see an elephant, and that we are certain that this visual perception of the elephant is not illusory, then we do not normally go on to infer that the yonder animal that we are seeing is an elephant, since it is making a trumpeting sound that can be made only by elephants. In such cases an attempt to establish the fact that the said creature is an elephant would be vitiated by the defect that is technically known as ***siddhasadhana*** i.e. establishment of what has already been established previously. The second situation where an inference becomes inoperative is where the absence of a certain entity is already established in the locus where that very entity is sought to be proved by some inference. In such cases, the inference in question is vitiated by the defect known as

[1] Diacritical marks not given

paramanantarabadhitatva, *i.e*. being contradicted by some other accredited means of knowledge.

With these prefatory remarks, we may now proceed to discuss the disputes that took place among some of the eminent philosophers of ancient India regarding the nature of ***anumeya***, *i.e*. what is sought to be established by inference.

II

We have said above that inference (***anumana***) has been admitted as a source of knowledge by many schools of Indian Philosophy. We have deliberately refrained from using the expression "*all* schools of Indian Philosophy", because there were some thinkers who questioned the very possibility of obtaining knowledge from inference. Some of them maintained that one can never establish with certainty the pervasion (***vyapti***) between the ***hetu*** and the ***sadhya*** that feature in the inference. The claim that whatever is characterised by the ***hetu*** is also characterised by the ***sadhya*** can be established beyond reasonable doubt only when we know *all* the instances of the ***hetu*** and all the instances of ***sadhya***; and this is feasible for only those who are omniscient. We ordinary mortals may observe very many cases of ***hetu*** and ***sadhya***, but even if all these observations reveal the co-existence of the ***hetu*** and the ***sadhya*,** they cannot establish conclusively that there are no counter-examples to the pervasion, *i.e*. cases where the ***hetu*** may exist even in the absence of the ***sadhya***. If the very basis of inference be open to doubt, then how can we claim that inference is an infallible source of knowledge?

Some other critics have tried to reject ***anumana*** as a source of knowledge by showing that it is impossible to specify the exact nature of **anumeya**, i.e. the object that is supposed to be established by inference. Now, if there is *nothing* to be established by inference, then we must give up the claim that inference is a source of knowledge. While developing this argument, they have considered the different permutations and combinations of the possible types of ***paksa*, *sadhya*** and ***hetu*** for showing that none of these possible combinations can be justifiably treated as a case of ***anumeya*** proper. If this claim can be logically established, then ***anumana*** will certainly cease to be an accredited source of knowledge.

The philosophers who admit ***anumana*** as a source of knowledge had tried their level best to meet such objections. Their counter arguments against the first objection mentioned above are well known, and need not be discussed here. But their counter replies to the second objection stated above are not so well known; and hence, some discussion of them will perhaps not be out of place. In order to discuss such counter replies, we have to take into account the various alternatives regarding ***anumeya*** that the critics of ***anumana*** have formulated, examine whether the list of such alternatives is really exhaustive, and then see whether any one of these alternatives can withstand the arguments advanced against it by critics.

III

Let us once again take a look at the full-blown form of the stock example of inference that is supposed to be employed for establishing the existence of fire in a hill on the basis of smoke that is present in that hill:

1) Yonder hill is characterised by fire.
2) Because there is smoke in it.
3) Wherever smoke is present fire is present, i.e. a kitchen.
4) The hill is also like that (i.e. characterised by smoke).
5) Therefore, the hill is also characterised by fire.

In this inference, the hill is the ***paksa*****,** fire is ***sadhya*** and smoke is ***hetu***. Since the hill is characterised by fire and smoke, the latter are called ***dharma***-s (i.e. the properties or characteristics) of the hill; while the hill is said to be the ***dharmi*** (i.e. the property-bearer or that which is characterised). Apparently, what this inference aims at proving or establishing has been stated in sentence no.1 stated above, while sentence no.5 claims that what was supposed to be established by this inference has actually been established. The critics of ***anumana*** now ask – what is the exact ***anumeya*** that is supposed to be proved by this inference? In the sentence 'the hill is characterised by fire', there are at least three elements – i) the hill (which is the ***dharmi***), ii) the fire (which is the ***dharma***), and iii) a certain relation (***sambandha***) that obtains between the fire and the hill, in the absence of which no characterisation could take place between the hill and the fire. When a certain relation R obtains between two entities X and Y, either X is qualified by Y, or Y is characterised by X. Accordingly, there are the following entities which could possibly be the ***anumeya*** of this inference: i) only the ***dharmi***, ii) only the ***dharma***, iii) both the ***dharmi*** and the ***dharma****,* being considered only by themselves, iv) the ***sambandha***, v) the ***dharmi*** qualified by the ***dharma*** (***dharmavisistadharmi***) vi) the ***dharma*** qualified by the ***dharmi*** (***dharmivisistadharma***). The critics of ***anumana*** now claim that none of these alternatives can be said to be ***anumeya*** here, and in the absence of any other alternative, it follows conclusively that there is nothing that can possibly be established by the influence that has been mentioned here. Similar arguments can be given in the case of other inferences, and this shows that no ***anumana*** or inference can establish anything whatsoever.

It is not known for certain who these critics of ***anumana*** are, and their works have not come down to us; but the arguments formulated by them have been noted and discussed by many early Indian philosophers who *have* admitted ***anumana*** as a source of knowledge. Among them, the names Dinnaga (a famous Buddhist thinker and the author of books like ***Pramanasamuccaya***, ***Hetucakra***, ***Nyayamukha*** *etc*.), Kumarila Bhatta (a follower of the ***Purva-Mimamsa*** school and author of books like ***Slokavarttika***, ***Tantravarttika***, *etc*.), Jayanta Bhatta (a follower of the ***Nyaya*** school and the author of ***Nyayamanjari***) and Cakradhara (the author of ***Nyayamanjari-granthibhanga***, a commentary on Jayanta's book) deserve special mention. The original works of Dinnaga are mostly lost, though translations of some of them in Chinese and Tibetan are available. We have, however, at our disposal many verses of Dinnaga that have been quoted in ***Nyayavarttikatatparyatika*** of Vacaspati Misra, another noted adherent of ***Nyaya*** school.

Vacaspati has quoted the following verses of Dinnaga where the objections of these critics and Dinnaga's answer to them have been given in a condensed form:

kacid dharmantaram meyam lingasyavyabhicaratah /
sambandham kecidicchanti siddhatvad dharma-dharminoh //
lingam dharme prasiddham cet kimanyat tena miyate /
atha dharmini tasyaiva kimartham nanumeyata //
sambandhe'pi dvayam nasti sasthi sruyeta tadvati /
avacyo'rthagrhitatvanna casau lingasangatah //
lingasyavyabhicarastu dharmenanyatra darsyate /
tatra prasiddham tadyuktam dharminam gamayisyate //

[***Nyayavarttikatatparyatika***, pp.147-148]

These verses have been briefly explained by Vardhamana Upadhyaya in his ***Nyayanibandhaprakasa***, and we have depended on this explanation for the exposition of the arguments stated in these verses. [***Nyayanibandhaprakasa***, p. 748]

Let us now see on what grounds these critics maintain that none of the alternatives listed above can be considered to be the ***anumeya*** proper of this inference. We start here by noting the arguments as recorded by Dinnaga in his ***Paramanasamuccaya***. That the dharma (*i.e.* some fire) is existent somewhere is already known to us through perception etc., and hence, an attempt to establish it once again by inference would involve the defect known as ***siddhasadhana***. The same is true of the ***dharmi*** (*i.e.* the hill), which is also known to us with certainty prior to the employment of this inference. Nor can **dharma** and ***dharmi*** taken together be the ***anumaya***, since both of them have been known to exist before the employment of this inference, and employment of inference for establishing them once again would obviously involve ***siddhasadhana***. It may now be claimed that even though the hill and some fire may be known previously, the fact that in this particular case, the hill is *related* to some fire was not known prior to the employment of this inference; and hence, this relation may be said to be the ***anumeya***. But such an answer is not tenable for various reasons. An ***anumeya*** is established on the pervasion that obtains between the ***anumeya*** and the ***hetu*** (which happens to be the smoke in this case). Now, such a pervasion does not obtain between the smoke on the one hand and the relation between the hill and fire on the other. No one can claim that wherever and whenever there is smoke, there is also the relation between the hill and fire. Besides, had this relation been the ***anumeya*** proper in this case, the first sentence stating the thesis of this inference should have been of the form 'there is some relation between the hill and fire', which is not the case. Moreover, upon hearing the sentence 'the hill is characterised by fire on account of smoke', we know *by implication* that there is some relation between the hill and the fire. Thus, the said relation, which is grasped from the cognition generated by this sentence, need not be established again by employing an inference. The last two alternatives are also ruled out by the fact that smoke is not pervaded either by (i) the hill characterised by the fire, or (ii) the fire characterised by the hill; and hence, neither the hill characterised by the fire nor the fire characterised by the hill can be established by smoke. Consequently, none of these two alternatives can be admitted here as the required ***anumeya***.

Kumarila Bhatta and Jayanta Bhatta have recorded some other arguments that were employed for rejecting the suggestion that here, the anumeya is ***dharmivisistadharma***, which in this case would be fire qualified by hill (*i.e.* the fire located in hill). According to Kumarila Bhatta, the critics have shown that one can think of eight alternatives in this case, which are as follows:

(i) Fire as such qualified by some location as such is the ***anumeya***,

(ii) The specific fire that was apprehended during apprehending the pervasion between smoke and fire, and qualified by some unspecified location, is the ***anumeya***,

(iii) The specific fire mentioned in alternative (ii) that is qualified by all locations, is the ***anumeya***,

(iv) Fire as such (*i.e*. some unspecified fire) qualified by the location that is proximate (*eg*. the hill) is the ***anumeya***,

(v) The specific fire mentioned in alternative (iv) that is qualified by the location that is proximate is the ***anumeya***,

(vi) The specific fire that is located in the hill, which is qualified by some unspecified location is the ***anumeya***,

(vii) The specific fire that is located in the hill, which is qualified by the kitchen (where the pervasion between smoke and fire was observed) is the ***anumeya***,

(viii) The specific fire located in the hill that is qualified by the hill is the ***anumeya***,

A little reflection will show that some of these alternatives [*eg*. alternative (i)] is vitiated by the defect known as ***siddhasadhana***. Any fire whatsoever must be located at some place or other – this need not be established by employing an inference. The other alternatives are obviously contrary to facts [*eg*. alternative (vii)]. Since all these alternatives are untenable, ***dharmivisistadharma*** cannot be admitted as the ***anumeya*** here.

[Kumarila Bhatta has discussed this issue in Verse nos.39-46 of the ***Anumanapariccheda*** of his ***Slokavarttika***. We have taken the help of the commentaries of Umbeka Bhatta and Parthasarathi Misra on these verses.]

It is not at all clear on what basis these alternatives (enumerated by Kumarila Bhatta) have been formulated; nor is it quite evident at first that the list of these alternatives is exhaustive. This, however is not the case with the list noted by Jayanta Bhatta, who has noted here *sixteen* alternatives; because the principle according to which these sixteen alternatives have been enumerated has been stated clearly in a verse quoted by Cakradhara in his ***Nyayamanjari-granthibhanga***, (which is a commentary on Jayanta Bhatta's ***Nyayamanjari***), and the way in which these alternatives have been framed also ensures the exhaustiveness of the list. The said verse reads as follows:

sarvonirdharitah purvah sailastho'gniscaturvidhah /
pratyekam sadhyate sarvapurvanirdharitadribhih //

[***Nyayamanjarigranthibhanga***,p.61]

According to this verse the ***dharma*** (i.e. the fire) here may be either of the four types, *viz*. (i) all fires, (ii) some unspecified fire, (iii) the fire that was apprehended earlier (in the kitchen while grasping the pervasion between fire and smoke), and (iv)the fire that is located on the hill. Likewise, the ***dharmi*** here may be either of the four types, *viz*. (i)all locations ,(ii) some unspecified location, (iii) the location apprehended previously (*eg*. the kitchen where the pervasion between the smoke and fire was apprehended earlier), and (iv) the hill. Now, each of the four types of the fire may logically be related to (or qualified by) any one of the four types of locations. Thus we have 4x4=16 possibilities, and as far as we can see, no further alternatives can be considered here, which shows that the list of alternatives is exhaustive. The full list of these sixteen alternatives about ***anumeya*** is as follows;

(i) All fires that are present in all locations.

(ii) All fires that are present in some unspecified location.

(iii) All fires that are present in previously perceived locations (*eg*. kitchen, *etc*.).

(iv) All fires that are located in locations like the hill, that are now being perceived.

(v) Some unspecified fire that is present in all locations.

(vi) Some unspecified fire that is present in some unspecified locations.

(vii) Some unspecified fire that is present in some previously perceived locations.

(viii)Some unspecified fire that are present in locations being perceived.

(ix) The fire that was perceived earlier (at the time of apprehending the pervasion of smoke by fire) that is present in all locations.

(x) The fire that was perceived earlier, and which is present in some unspecified location.

(xi) The fire that was perceived earlier, and which is present in some previously perceived location.

(xii) The fire that was perceived earlier, and which is present in some location that is being perceived.

(xiii) The fire on the hill (which is sought to be established) and which is present in every location.

(xiv) The fire on the hill that is present in some unperceived location.

(xv) The fire on the hill that is present in some previously perceived location.

(xvi) The fire on the hill that is present in some location that is being perceived (*eg*. the hill).

Here again, some of the alternatives are vitiated by ***siddhasadhana***, while the others are contradicted by perception, or are self-contradictory. This, according to the critics, shows that ***anumana*** cannot be a ***pramana***.

IV

Dinnaga, Kumarila Bhatta and Jayanta Bhatta, having considered such dialectical arguments against the possibility of any ***anumeya***, try to support the view that ***anumeya*** is ***dharmavisistadharmi***. We have already_noted that the critics of ***anumana*** have rejected this view on the ground that since the ***hetu*** (*eg*. smoke) is not pervaded by ***dharmavisistadharmi*** (*eg*. the hill characterised by fire), the former cannot establish the latter. Dinnaga, however, seeks to rebut this argument by pointing out that if it so happens that some ***hetu***, which is itself established in some ***dharmi***, is capable of establishing in that very ***dharmi*** some ***dharma*** that pervades that particular ***hetu***. Thus, if it is known with certainty that smoke is present in a hill, and if it is also known for certain that smoke is pervaded by fire, then that smoke is capable of establishing the existence of fire in that hill.

This is the import of the sentence "***tatra prasiddham tadyuktam dharminam gamayisyate***". Vardhamana Upadhyaya's comments on this line are –"***tatrapakse, (pra)siddham lingam tadyuktam agniyuktam dharminam parvatam sadhayisyati – ityarthah***". Kumarila Bhatta also expressed a similar view in verse no.47 of ***Anumanapariccheda*** of his ***Slokavarttika***:

tasmad dharmavisistasya dharminah syat prameyta /
sa desasyagniyuktasya dhumasyanyaisca kalpita //

He has indicated this view even earlier in the same chapter:

sa eva cobhayatmayam gamyo gamaka eva ca //
asiddhenaikadesena gamyah siddhena bodhakah /

[***Anumanapariccheda***, verses 24b-25a]

Parthasarathi Misra explains this by saying that the same hill, in so far as it is to be characterised by a property that was not established earlier (*viz*. Possession of fire) is the ***anumeya*** here, whereas that very hill, which is characterised by the smoke (which has been established in the hill, and which is also pervaded by fire) functions here as the ***gamaka***, *i.e*. something that can establish the ***anumeya***:

asiddhena pramanantaraprasiddhena vahnimattvatmana'numeyatvam, pramanantarasiddhena tu dhumavattvatmana gamakatvam iti.

[***Nyayaratnakara***, p.252]

Jayanta Bhatta has also supported this view in an almost identical manner:

tasmadagnivisistah paridrsyamano desa eva
sadhyah. sa ca svarupata pratyakso'pi
paroksadharmavisistataya anumeya iti
dhumadharmayogat agnidharmavan sa evanumiyate.
so'numanasya visayo lingityucyate.
paroksagrahanam caitadabhiprayena.

[***Nyayamanjari***, Vol.I,p.311]

The view of Dinnaga, Kumarila Bhatta and Jayanta Bhatta regarding ***anumeya*** was not unanimously adopted by other Indian philosophers who have admitted ***anumana*** as a source of knowledge. Thus, for example, Vatsyayana, author of ***Nyayabhasya***, while commenting on ***Nyayasutra*** 1.1.5 has said that in the case under consideration, fire is the ***anumeya*** that is established here by smoke ("***yatha – dhumenagniriti***". ***Nyayabhasya***, p.). Uddyotakara, the author of ***Nyayavarttika*** (which is a commentary on ***Nyayabhasya***), differs from Vatsyayana and maintains that in the case under consideration, what is established is the fire that is the characteristic of smoke of a certain kind. Thus, for Uddyotakara, the ***paksa*** and **hetu** here are not the hill and smoke, respectively, but smoke and certain properties of smoke (*eg*. continuous and upward movement from a certain place), from which it can be known that a particular smoke having those properties is accompanied by fire. [See ***Nyayavarttika***, p.]

This view of Uddyotakara (which has been mentioned by Kumarila Bhatta in the phrase "***dhumasyanyaisca kalpita***"), has not found favour with the later philosophers; most of whom agree here with Vatsyayana, and maintain that in the case under consideration, the ***anumeya*** or ***sadhya*** is fire. A detailed account of such views and arguments in favour of them would considerably increase the length of this paper; and hence, is omitted here. The interested reader may consult ***Nyayadarsana*** of Phanibhusana Tarkavagisa, Vol.I. pp.164-168 for an extremely scholarly and illuminating discussion of these views.

References

1. ***Nyayabhasya*** of Vatsyayana edited by Anantalal Thakur, Indian Council of Philosophical Research, New Delhi, 1997.
2. ***Nyayadarsana*** (Bengali translation of ***Nyayabhasya*** with elaborate exposition) by Phanibhusana Tarkavagisa, Vol.I (Reprint), Paschimbanga Rajya Pustak Parishad, Kolkata, 1981.
3. ***Nyayamanjari*** by Jayanta Bhatta edited with the editors own commentary ***Nyayasaurabha*** by K.S. Vardacharya, Oriental Research Institute, Mysore, Vol.1, 1969.
4. ***Nyayamanjarigranthibhanga*** by Cakradhara edited by Nagin J. Shah, L.D. Institute of Indology, Ahmedabad, 1972.
5. ***Nyayanibandhaprakasa*** by Vardhamana Upadhyaya edited (along with ***Nyayavarttikatatparyaparisuddhi*** of Udayana upto ***Nyayasutra*** 1.1.5) by V.P. Dvivedi and L.S. Dravid, Asiatic Society, Calcutta, 1911-1924.

6 ***Nyayavarttika*** of Uddyotakara edited by Anantalal Thakur, Indian Council of Philosophical Research, New Delhi, 1997.

7. ***Nyayavarttikatatparyatika*** of Vacaspati Misra edited by Anantalal Thakur, Indian Council of Philosophical Research, New Delhi, 1997.
8. ***Slokavarttika*** of Kumarila Bhatta with ***Nyayaratnakara*** of Parthasarathi Misra edited by Swami Dwarikadas Shastri, Tara Publications, Varanasi, 1978.
9. ***Slokavarttika*** of Kumarila Bhatta with ***Tatparyatika*** of Umveka Bhatta edited by S.K. Ramanatha Sastri, Revised by K. Kunjunni Raja and R. Thangaswamy, University of Madras, second revised edition, 1971.

Prabal Kumar Sen
Department of Philosophy
University of Calcutta
pksen_cu@rediffmail.com

Some Musings on Mathematical Creations and Rationality

Smita Sirker

Abstract. Chakraborty emphasises on the creator-discoverer role of a mathematician. In his writings he tries to outline a picture of mathematics which represents the beauty of the discipline on one hand, and on the other, the rigorous turmoil, often painful, that an artist-mathematician has to undergo to mother a mathematical creation. Judging the rationality of the mathematical pursuit or the created artefact of an artist-mathematician is no simple task. His writings resonates the idea that a pregnant mathematical thought is nurtured within a culture, within a particular socio-political milieu and from a given historical lineage. This article engages in some reflections on Professor Chakraborty's view of an 'artist' in a 'mathematician' and the rationality involved in mathematical pursuits.

1. The Beginning

My initial conversations with Professor Chakraborty were fascinating and yet very surprising. It took some time to overcome my surprise before beginning to understand his, I would say, unique perspective on the nature of a mathematician's enterprise, task and effort. I have never heard any mathematician speak like this before (not that I know many of them) and I may add, this was the first time I could comprehend what a 'mathematician' was actually saying. He gave me the opportunity to 'speak to' and 'speak with'[1] a mathematician/logician. Chakraborty says in 'Sundarer Satya, Ganiter Satya'[2] that ordinarily 'mathematics' means 'computation'/'calculation' and we grow up learning and mastering only the process of computation, thereby gradually moving away from the essence of the subject and its historical development. Knowing mathematics simmers down to mastering the skill of computation or failing to do so. It never remains as a matter of 'fondness' for the beauty of mathematics itself. My

[1] This distinction of 'speaking to' and speaking with' is maintained by Professor Shefali Moitra. 'Speaking to' is a mode of communication like a 'form of transmission', which is a necessary but not a sufficient condition of communication. 'Speaking with' is a mode of communication which involves 'participation' (a kind of a co-operative effort) over and above 'transmission'.
Moitra, S.: 'Speaking to…' and 'Speaking with…'. In: Feminist Thought: Androcentrism, Communication and Objectivity. Munshiram Manoharlal Publishers Pvt. Ltd., New Delhi (2002) 81-99.

[2] A rough translation of this Bengali title would be 'Truth of Beauty, Truth of Mathematics' in Ganiter Dharapat O Galpasalpa (in bengali). 5th edn. Nandimukh Samsad, Kolkata (2011)

initial learning of mathematics was typically confined to mastering the process of ‘doing sums’ which I never really liked. As mathematical problems got complex my interest also faded beyond revival. Never had I thought in my wildest dreams that I could meaningfully ‘open a dialogue with a mathematician (logician)’. So writing this article and contributing to this volume is still not believable.

I was working on the question of ‘human rationality’ – the challenges the concept poses to decipher many of our activities as rational or otherwise. Starting from the question of what is the ‘nature of rationality’; there are a host of problems which surfaces when we try to evaluate many human actions based on any defined standards of rationality. Professor Chakraborty pointed out that one such human activity was ‘a mathematician’s pursuit of a mathematical problem’. This was our area of convergence and we started thinking on - how to ‘rationalise’ or ‘de-rationalise’[3] some of the long-standing mathematical pursuits. A unique dimension to this problem was given by Professor Chakraborty’s another perspective and firm belief that mathematicians are *artists* and their work (‘creations’) are *objet d'art*. Two questions evolved out of our discussions: (i) Are all mathematical pursuits ‘rational’? and (ii) Is the search for a ‘rationality’ dimension mandatory for every mathematical artefact? These questions are very tricky and demand very lengthy discussions and unfolding of many intricate, intertwined layers of rationality, objectivity, criterion of universal acceptability and the rationale of subjective pursuits. We are still working towards drawing meaningful answers to these questions. Given a brief backdrop of our research orientation, my primary aim of this article will be to present some musings on Professor Chakraborty’s view of an ‘artist’ in a ‘mathematician’ and what could be his possible rejoinder to the rationality facet of a mathematical creation.

2. Mathematical Objects: *Existence By Description*

> গণিতশিল্পী ছাড়া ‘আকাশকুসুম’ -এর চাষ-আবাদ করার ক্ষমতা ও দুঃসাহস বিজ্ঞানের জগতের আর কোন সাধকের থাকতে পারে? বিজ্ঞান জগতের বাইরে (সত্যিই বাইরে কি?) হয়তো আছে কবির, হয়তো শিল্পীর।
>
> \- Mihir Kr. Chakraborty[4]

---

[3] I take the liberty to introduce the phrase ‘de-rationalise’ with a particular intention in mind. The question is ‘Can we ‘de-rationalise’ any mathematical activity if it fails to meet the expectations of the mathematical community?’ Chakraborty (see Chakraborty, M. K.: Mathematical RUPAs and their Artists. In: PHISPC-CONSSAVY, vol13, Part 6, Centre for Studies in Civilizations, New Delhi (accepted for publication)) maintains that there is no single mathematical community – there are many. This ‘multiplicity’ and ‘multi-directedness’ in mathematical community is opportune to let grow of many weeds which may be deviant of the mainstream expectations. But these weeds have their own beauty and joy of creation and expansion. Hence how do we deny rationality to such pursuits? Is it justified to do so?

[4] Akashkusumer Adhikar (trans. Right to Imaginaries). In: Ganiter Dharapat O Galposalpo. 15

Chakraborty in his writings tries to outline a picture of mathematics which represents the beauty of the discipline on one hand, and on the other, the rigorous turmoil (often painful) that a mathematician has to undergo to mother a mathematical creation. This 'pain' is both literal and metaphorical. It is literal as it expresses the toil that a mathematician has to undertake to give shape to his creation and present it with a distinctness of a created object. It is metaphorical in the sense that many mathematicians had to devote their entire life to resolve long-pending, knotty mathematical problems. Such instances are plenty. One just needs to look back in time and see the labour invested by great mathematicians to prove Euclid's fifth postulate and the gradual development and hard acceptance of an alternate, non-Euclidean geometry. Mathematician Farkas Bolyai pleaded his son János Bolyai to give up his pursuit on Euclid's parallel postulate. He wrote to his son, "For God's sake, please give it up. Fear it no less that the sensual passions because it, too, may take up all your time and deprive you of your health, peace of mind and happiness of life."[5]

Questions regarding the role of a mathematician have been raised by many. Do mathematicians discover the 'truths' that are already there in the world or do they create the mathematical truths in their mental laboratory? Are mathematicians' discoverers or are they creators? According to Chakraborty, a mathematician can be both, though there is an order between the two roles. First, a mathematician is a creator and then a discoverer. He says, "Is not an artist a discoverer too? He creates but this creating is also a discovering. From the very first stroke on the blank canvas he enters into a dialogue with it. The canvas starts demanding, the artist may or may not satisfy the demands, may or may not be able to discover the 'needs' that are radiated from the canvas in its process of becoming an artistic creation."[6] Let us unfold his view.

According to Chakraborty, the objects of investigation for a physicist or a biologist exist independently of the knower or the investigator. The truths about these objects exist prior to the findings of the physicist or the biologist unveiling their research results. The truths exist even before the linguistic explanation of the physicist or the biologist and the truths continue to exist without the knower. However, the mathematical objects are not out there in the physical world. Mathematical objects are there in the minds of the mathematicians, in their imagination. Mathematicians need an expression, a language to make their objects public. He says, for others, like physicists, biologists, botanists, etc. it is '*description after existence*' but for the mathematician it is '*existence by description*'.[7] Thus for a mathematician, engaging

[5] Davis J. P. & Hersh, R.: The Mathematical Experience. Birkhäuser, Boston (1981). 220-221
[6] Mathematical RUPAs and their Artists. 4
[7] Akashkusumer Adhikar. 17

in a linguistic explanation and description (mathematical proof)[8] is not a meta-level activity, unlike the physicists, biologists, etc. To quote Chakraborty:

> "A mathematician enters into the act of creativity, consciously or sometimes semi-consciously too and a mathematical object is created. It is a piece of artefact produced by the mathematical 'artist' or 'artisan' if one prefers to call her or him. Others have their 'objects' of study out there, outside the knower, independently existent. All these objects remained even before the utterance of a single sentence about them by humans. Mathematical sciences or *arts* [emphasis mine], on the other hand, have to make some initial claims in order to create mathematical objects which, like art-objects, are constructions of mind, the mind of the mathematician."[9]

Chakraborty emphasises on the creator-discoverer role of a mathematician. According to him, after the mathematical objects are created in the mind's eye of the mathematician, he engages in the process of discovering the properties of his created object. The creation of a mathematician is like that of an artist's painting on a blank canvas which starts with a few simple strokes and then transforms into an elaborate art-work – through exchanges between various forms of imaginations.[10] The abstract mathematical objects (RUPA's[11] as Chakraborty calls it) created by the mathematician carries with it certain properties unknown to its creator. These properties are the true stories of the mathematical constructs. The discovery and unveiling of these properties is a second, meta-level activity that the mathematician creator has to engage in. For this discovery, there is a need for a laboratory setup where the primary instrument is logic and the process is mathematical proof.[12] For any dweller of the mathematical world, the task is twofold. As an *artist*, a mathematician creates extraordinary, unique, abstract, exotic mathematical RUPAs. And as a *discoverer*, a mathematician explores the unknown properties of the created mathematical RUPAs.

> "The new RUPA appears alien, mysterious and elusive to its creator. The creator has ascribed in it some initial properties and a few more properties might also be visible – but what else are there in it hidden? Unravelling its mystery, then, becomes a compelling, unavoidable task to the mathematical artist. She enters into a dialogue with the mathematical object she has created just like a natural scientist enters into a dialogue with nature."[13]

---

[8] More appropriately, the mathematician engages in a language-game with oneself while being engaged in the process.

[9] Mathematical RUPAs and their Artists. 1

[10] Ibid.

[11] It is a Bengali word whose meaning can be expressed as 'the beauty (usually visual) of a form along with its aura'. The word expresses the beauty that the object beholds in the eye of the observer.

[12] Akashkusumer Adhikar. 18

[13] Mathematical RUPAs and their Artists. 4

3. Visualization Of A Mathematician

> "The dialogue is a game – a game of responding to each other. In case of natural scientist the other player is nature, in case of an artist it is the canvas while in case of mathematical-artist the other partner is the object created by herself. .. She sets the rules for reasoning and waits for the answers. .. The process of getting a response from a mathematical object is like locating the pains in a mute patient. Answers are obtained through mathematical proofs. .. A proof is a conversation between the mathematician and her created object – this object being ultimately laid down by a few axioms – a few initial claims about the object or a set of similar objects."
>
> - Mihir Kr. Chakraborty[14]

Chakraborty suggests that mathematical proofs are instrumental in discovering the properties of a created mathematical object. It is through the process involved in the development of a mathematical proof that a mathematician enters into a conversation with his own creation. He would say that there is also an element of beauty in mathematical proofs – which are often termed as 'elegant proofs'. Such an identity is reflective of an admiration which is often held for creative arts. That is the reason why Chakraborty finds resemblance between creations of musical pieces and creations of mathematics. Such a visualization of mathematical objects is quite contrary to a Platonian approach.

Plato used *eidos* to mean the inner structure or intelligible form of a thing, the shape grasped by the eye of the soul, which serves as the object of knowledge or thought.[15] Forms exist in some way independently of particulars, and the latter are imperfect copies of the Forms. Thus Plato gives us the *Two World Theory* – the world of particulars and the world of eternal Forms. According to him, the eternal Forms of mathematical objects exist in the world of Forms. Brown enumerates some features that constitute the core of Platonism, following the thoughts of Frege, Hardy, Gödel, who were Platonists in many ways. To list a few:[16]

i. Mathematical objects are perfectly real and exist independently of us. We do not create them in any way, rather we discover them.
ii. Mathematical entities are abstract in one sense, but not in another. Abstract stands for two distinct meanings. One sense pertains to *universals* and *particulars*. Another sense pertains to *outside space and time*, not concrete, not physical.
iii. We can intuit mathematical objects and grasp mathematical truths.

---

[14] Ibid., 4-5

[15] Bunin, N., & Yu, J.: The Blackwell Dictionary of Western Philosophy. Blackwell Publishing Ltd, UK, 2004, p. 264

[16] Brown, Robert, J.: Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures. Routlegde, London (2002) 11-14

Frege distinguishes between (a) *ideas* (they are psychological entities); (b) *thoughts* (they are content of our *ideas*); and (c) *sentences* we use to express them (they are *things of the outer world*, like trees, electrons, sound waves). His *thoughts* are Platonic entities. Frege says, "So the result seems to be: thoughts are neither things of the outer world nor ideas. A third realm must be recognized. What belongs to this corresponds with ideas, in that it cannot be perceived by the senses, but with things, in that it needs no bearer to the contents of whose consciousness to belong. Thus the thought, for example, which we express in the Pythagorean theorem is timelessly true, true independently of whether anyone takes it to be true."[17] Hardy presents an interesting way of looking at developing mathematical proofs. For him, mathematicians engage in acts of discovery. "I have myself always thought of a mathematician as in the first instance an *observer*, a man who gazes at a distant range of mountains and notes down his observations. His object is simply to distinguish clearly and notify to others as many different peaks as he can. .. there is, strictly, no such thing as mathematical proof; that we can, in the last analysis, do nothing but *point*; that proofs are what Littlewood and I call *gas*, rhetorical flourishes designed to affect psychology, pictures on the board in the lecture, devices to stimulate the imagination of pupils."[18]

Chakraborty would be less sympathetic to the above visualisation of mathematics, though he will not be dismissive. I guess that he will not endorse a reductionist stance concerning the way mathematics should be visualised and pursued. His heart would still go out to distinguish mathematics as an artistic game, where the artist-mathematician is engrossed in the madness of creation. Doing mathematics is like modelling a lump of clay which is non-definitive at first, goes through constant re-modelling until the mathematician sees a well-formed structure to end the game of his creation.

4. The Private Game Of Creation And The Public Game Of Communication

Let us enter into the private-public domain of mathematical ventures. Chakraborty remarks "Once produced, the product becomes public. Each individual subject in the public who encounters the product sets up a relationship with it – a relationship that is basically private."[19] This position is extremely intriguing and as well challenging. We will consider two sharp minds, Brouwer and Wittgenstein, to address the private-public debate.

Brouwer proposal for a foundation of mathematics led him to *intuitionism*, which viewed mathematics as a free activity of the mind, independent of any language or

[17] Frege, G.: The Thought: A Logical Inquiry. Mind. 65: 259 (1956) 302
[18] Hardy, G. H.: Mathematical Proof. Mind. 38:149 (1929) 18
[19] Mathematical RUPAs and their Artists. 1

Platonistic interpretation of mathematical concepts and objects. Mathematics was a free activity of exact thinking founded on pure intuition. Mathematical truths are not dependent or determined by any objective facts of the world. A proposition only becomes true when the subject or the thinker has experienced its truth by constructing an appropriate mental construction. The falsity of any proposition, likewise, is also experienced by the subject when it is realised that an appropriate mental construction is not possible. Brouwer's understanding of the mathematical activity is *private*, almost solipsistic. Mathematicians do not discover pre-existent things. They engage in a languageless mental activity. He says:

> "The first act of intuitionism completely separates mathematics from mathematical language, in particular from the phenomena of language which are described by theoretical logic, and recognizes that intuitionist mathematics is an essentially languageless activity of the mind having its origin in the perception of a move of time, i.e. of the falling apart of a life moment into two distinct things, one of which gives way to the other, but is retained in memory. If the two-ity thus born is divested of all quality, there remains the empty form of the common substratum of all two-ities. It is this common substratum, this empty form, which is the basic intuition of mathematics."[20]

Though Chakraborty endorses a private realm in mathematical creation, he will not consider the process to be a languageless activity. In fact, without language the task becomes difficult if not impossible. Mathematical existence for him is *existence by description*. Physicists and biologists, who have their objects of study at their disposal, description becomes a second-level or a meta-level activity. But for the mathematicians there are no such visible objects that await description. Mathematical objects and creations are non-communicable without language. Thus though the activity is primarily private, it is not essentially *only* private and thereby solipsistic.

Let us here look at what Wittgenstein had to say about private language. Mathematical activity for him is essentially *public* and the role of language in mathematical operations is important. Wittgenstein would argue against *private language*. Kripke says that Wittgenstein's main argument against the existence of private language is that "there is no fact of the matter about whether I am using a rule when I am considered individualistically; it is only my membership within a linguistic community that gives content to the notion of my being justified in using words as I do."[21] We use and share a language which is public, and which constitutes of rules developed by the community. According to Wittgenstein, the role of language in mathematical operation is crucial, as he thinks that to do a mathematical operation is to follow the rules laid down for that operation. Pradhan observes, "… to make a mathematical move is equivalent to making a move in the language according to rules. That is why Wittgenstein shows that a mathematical proposition has its use in

[20] Brouwer, L. E. J.: Historical Background, Principles and Methods of Intuitionism. South African Journal of Science. 59 (1952) 141-142

[21] Kripke, S.: Wittgenstein on Rules and Private Language. Harvard University Press, MA, Cambridge (1982) 89

ordinary language is so far as it leads to some new proposition in language. Mathematics and language are thus related in such a way that to understand the one is to understand the other. This does not, means that mathematics is a mere verbal game. It is a substantive activity carried on in language."[22]

Chakraborty's position is curiously both closer and yet very far from Wittgenstein's position. Chakraborty had remarked that once the (mathematical) product is created, then it becomes public. Then each member in the public who stumbles upon the product sets up a private relationship. The transition is like from the private realm (the creator's mind) to the public domain (the creation is out there in the public to respond to this creation) and back again into the private realm (the mind of the appreciator). It is both *private* and *public*. So long as the mathematician is involved with his creation, it is *private*, where the creator engages in a private dialogue (with his creations) within oneself. It is not soliloquy. The mathematician enters into a conversation with the *object* that he is visualizing, which maintains an outward silence. However, there is also a parallel preparation to make it public (thus the language engaged in the creative process is not private in the absolute sense, i.e. non-shareable in essence). The process of creation is private, not the language. The moment the creation is ready to be presented it becomes *public*, and it is up to the community members to enter in to a relation with the creation. The third and often unspoken part is how a member enters into a relation with the creation – which Chakraborty holds to be a private one. His position is closer to Wittgenstein because language for both of them is an integral part of mathematical exercise (which stands in contrast to Brouwer's position). However, Chakraborty's position is far from Wittgenstein because mathematical creations start from a private domain, and if a mathematical creation takes a well-formed shape then only it enters into a public game of acceptance or rejection of the creation and the creator by the community.

5. The Rationality of Mathematical Pursuits: A Few Questions

One important question that Chakraborty has often raised is about the situatedness of a mathematician in one's given socio-cultural milieu. His thoughts resonates the idea that a pregnant mathematical thought is nurtured within a culture, within a particular socio-political milieu and from a given historical lineage. It is almost impossible to do justice in understanding the works of a mathematician without considering his *locatedness*. One's socio-political as well as cultural situatedness silently contribute to one's creativity. One cannot study mathematical inventions without its historical situatedness. Chakraborty reiterates the significance of studying the creator prior to his creations. I will conclude this article by ruminating on mathematical creations and rationality.

[22] Pradhan, R.C.: The Great Mirror: An Essay on Wittgenstein's Tractatus. Kalki Prakash, New Delhi (2002) 34

This issue has been a part of many of our discussions. How do we decide on the fruitfulness or worthiness of a mathematical pursuit? How do we rationalise the many (instances of) failed attempts of mathematics? If a mathematician pursues a long-standing, unresolved mathematical problem and eventually fails to resolve it – would someone relegate the mathematician's effort to the domain of irrationality? One might say that venturing into an attempted yet unresolved knotty problem would itself be an irrational move. Consider cases like mathematicians' struggle to prove Euclid's parallel postulate. Or will Gödel's incompleteness theorem render every mathematician's effort to establish completeness of a system valueless? What criterion of rationality is demanded of mathematical pursuits? In one of our discussions, I had asked him – *how would he see the question of rationality posed for mathematical creations?* While reflecting on this question, he raised many of the above questions which I just stated.

Judging the rationality of the artefact of an artist-mathematician is no simple task. Mathematical products as works of art can have its own intrinsic value. But mathematical communities are often not ready to accept mathematical creation at its mere face-value. It must serve a purpose which will determine its value. Hence the value of the creation is at the mercy of it functionality and use. Thus, a mathematical pursuit as a rational act is determined by the usefulness of the created product.

> তবে কি ব্যবহারযোগ্যতাই মাপকাঠি? গণিতের কোনো ধারণা বা তত্ত্বের সাফল্য কি নির্ভর করবে তা কি পরিমাণে অন্য ক্ষেত্রে কাজে আসে তার উপরে?[23]

However, Chakraborty asks us to also focus on the process of creation – which itself can be worthy. The joy of creation is no less than the product. Even a failed attempt can be worth experiencing. Do we judge whether a painter's painting or a poet's poetry to be rational? If a painter's work is not only unappreciated but also severely criticised – what would be the value of the painting? A painter might bask in the effort of the process of creation without minding much about the end result. Would mathematicians also take a similar stance as that of a painter? Or is there a difference between the two creations – one demands of worthiness (mathematics) and the other does not.

The question of rationality regarding mathematical creations is yet to be resolved at our end. But the question is alive and we agree that the process of finding an answer is itself stimulating.

References

1. Brouwer, L. E. J.: Historical Background, Principles and Methods of Intuitionism. South African Journal of Science. 59 (1952) 139-146

[23] Ganiter Dharapat O Galposolpo, 60

2. Brown, Robert J.: Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures. Routlegde, London (2002)
3. Bunin, N., & Yu, J.: The Blackwell Dictionary of Western Philosophy. Blackwell Publishing Ltd, United Kingdom (2004)
4. Chakraborty, M. K.: Mathematical RUPAs and their Artists. In: PHISPC-CONSSAVY, vol13, Part 6, Centre for Studies in Civilizations, New Delhi (accepted for publication)
5. Chakraborty, M. K.: Ganiter Dharapat O Galpasalpa (in bengali). 5th edn. Nandimukh Samsad, Kolkata (2011)
6. Davis J. P. & Hersh, R.: The Mathematical Experience. Birkhäuser, Boston (1981)
7. Diamond, C. (ed.): Wittgenstein's Lectures on Foundations of Mathematics, Cambridge 1939. Cornell University Press, New York (1976)
8. Diamond, C.: Wittgenstein, mathematics, and ethics: Resisting the attractions of realism. In: Sluga, H. & Stern, G. D. (eds.): The Cambridge Companion to Wittgenstein. Cambridge University Press, Cambridge (1996)
9. Dummett, M.: Wittgenstein's Philosophy of Mathematics. The Philosophical Review. 68:3 (1959) 324-348
10. Frege, G.: The Thought: A Logical Inquiry. Mind. 65:259 (1956) 289-311
11. Hardy, G. H.: Mathematical Proof. Mind. 38:149 (1929) 1-25
12. Hardy, G. H,: A Mathematician's Apology. University of Alberta Mathematical Sciences Society, Canada (2005) URL: http://www.math.ualberta.ca/mss/ (first electronic version 1.0)
13. Kripke, S.: Wittgenstein on Rules and Private Language. Harvard University Press, Mass, Cambridge (1982)
14. Moitra, S.: Feminist Thought: Androcentrism, Communication and Objectivity. Munshiram Manoharlal Publishers Pvt. Ltd., New Delhi (2002)
15. Nelson, E.: Mathematics and Faith. URL: http://www.math.princeton.edu/~nelson/papers.html
16. Pradhan, R. C.: The Great Mirror: An Essay on Wittgenstein's Tractatus. Kalki Prakash, New Delhi (2002)
17. Sirker, S., & Chakraborty, M.: Does Challenge To Truth And Consistency Also Redefine Boundaries Of Rationality?. Logos Architekton. 4:1-2 (2010) 7-27
18. Wrigley, M.: Wittgenstein's Philosophy of Mathematics. The Philosophical Quarterly. 27:106 (1977) 50-59

Smita Sirker
Department of Philosophy and
School of Cognitive Science,
Jadavpur University, Kolkata.
ssirker@gmail.com

Some pages from my diary

Sourav Tarafdar

After drinking tea and eating some biscuits, a man asked the poor lady shopkeeper "how much do I have to pay?" She replied "20 rupees". The man took out a 20 rupee note from his wallet and intended to give it, but suddenly he stopped. By noticing the note, which was not in a good condition because of its repeated change of hands, the man said "no no, I cannot give you this bad one" and then he looked into his wallet and took out the most fresh 20 rupee note he had and gave it to the shopkeeper with a smile which, for me always, has the power to hold me mesmerized. Yes, you are correct, the man was, Professor Mihir Kumar Chakraborty, my beloved Sir. On that day I was with him and I will never forget this incident.

On the first day of his course in my masters program in the University of Calcutta, I became attracted by Sir's teaching style. In his class he always gave some space to us to think by ourselves. After finishing graduation when we came to the masters course, I thought that Mathematics gives us the ultimate truths: once a theorem has been proved then there cannot exist any further scope of thinking about the theorem. But Sir was the first person who forced such thoughts of mine to undergo a change. I can remember that some times in his lecture, after proving something, he asked us "are you feeling okay with this whole thing? Is not something annoying you somewhere?" Before that I saw people asking these questions only to focus on the correctness of the proof, but Sir was the first person who showed us that we can ask such a question regarding the correctness of the statement of the theorem itself and not about the proof. There can be a huge debate on whether this kind of thinking is correct or not, but my point is, he showed us first that this kind of debates exists and that we should not let these things go without serious thinking. He told me "open your mind, don't take anything as granted."

I am attracted by the way Sir thinks and therefore, I am very lucky to get him as the supervisor of my research program on "Some Non-Classical Set Theories and Their Associated Logics."

Beside Mathematics I have lots of other things to learn from him. I am giving another instance here. Once in the lunch time of a seminar everyone was taking their food and discussing the topics of the seminar in groups. Sir was also in one such group. Suddenly I noticed that Sir was going a little bit further from the food court with his plate of food. I could not suppress my curiosity and followed him without saying anything. Then I saw there were five or six dogs and Sir was dividing the food in his plate in the same number of parts as there were dogs, and then serving it to them with the conversation, "come here, this one is for you", and then to another "you come here, this part is for you", and so on The amazing thing was that no dog was trying to eat another's food and the dogs were not quarrelling among themselves. At that very moment, some guards of that institute came with wooden sticks to beat the dogs, assuming that the dogs were disturbing a guest. But I saw that Sir started protesting and angrily said "why you are doing this? At night when you sleep who protects your institute?"!!! Both of us waited there till the dogs finished their lunch. The story does not end here. After finishing their food the dogs went away, but then Sir told me "Sourav, now we have to clean this place", and we cleaned the place as much as we could. This single five-minute incident taught me a lot. There are, I believe thousands of such things which I can remember, but the page limitation of this article does not allow me to narrate all of them.

When I went for the first time to Sir's house, the room where I sat was filled with lots of literature books and the total mass of them was so much that I could hide myself with some of my friends behind the heap of the books. I wondered that day that among them there was not a single book of mathematics (though in the other rooms there were lots of mathematics books, which I discovered when I visited Sir's house later). Beside Mathematics, he is very much involved with the modern arts and things which affect our lives and our very existence. This makes me wonder in astonishment about how he gets time to maintain all these things together with mathematics, and how does he manage time to accommodate so much of activity and involvement. I have noticed that sometimes he finds mathematics in pictures, and sometimes he finds poetry in mathematics. It seems that mathematics, poetry, literature, paintings... everything is tied together in a single thread in his life. May be this is why he can see mathematics from outside of mathematics.

There are never ending things which I can learn from Sir, but it is not true that I only want to learn something or other from him. On many occasions I feel free to share my joy, my sorrow, or my thoughts with him. I know that some of them may be meaningless or valueless to Sir. But he never comes up with such expressions, on the contrary he shows interest in these. It is for the fact that he bothers least about how knowledgeable or how famous he is, that he can mix with his students (rather, I should say, common people) by coming down to the level of the person he is talking to.

Sir is an idol to me. I have some photos of Sir in my laptop and sometimes I make these the wallpapers in my laptop. May be others will find this silly, but if a teacher can be such for his student then it is enough to reflect what the student feels about his teacher. I will consider my life a success if someday I can become only a tenth of what Sir is in my eyes.

Sourav Tarafder
University of Calcutta,
35, Ballygunge Circular Road, Kolkata
souravtarafdar@yahoo.co.in