Logic via Games

Rohit Parikh

rparikh@gc.cuny.edu

City University of New York

IIT Kanpur, January 2008

Two interpretations:

Let us consider two models $M =< N, <>$ and $M' =< Q, <>$ and two formulas, $A = (\forall x)(\exists y)(x < y)$ and $B = (\forall x)(\forall y)(x < y\ ra(\exists z)(x < z \wedge z < y))$

Now both A, B are true in M', and A is true in M as well, but B is false in M. Thus we have,

$M \models A$ but $M \not\models B$.

We now talk about the game interpretations of the two formulas. In this interpretation, we consider only formulas formed from literals using $\vee, \wedge, \forall, \exists$. Eloise controls $\vee, \exists$ and Abelard controls $\wedge, \forall$.

To be specific, we can associate two games G_A and G_B with A and B respectively.

In G_A, Abelard chooses an element a from the domain $\|M\|$, and Eloise picks an element b from $\|M\|$. If $a < b$ then Eloise wins. And clearly she has a winning strategy consisting of letting $b = a + 1$.

In G_B, since B begins with a universal quantifier, Abelard picks two elements a, b in $\|M\|$. We now have the formula $a < b \rightarrow (\exists z)(a < z \wedge z < b)$. i.e., $a \not< b \vee (\exists z)(a < z \wedge c < b)$

If $a \not< b$, then Eloise wins. If not, she has to pick a c. After that they are now at the formula $a < c \wedge c < b$. Abelard can pick either $a < c$ or $c < b$. If that formula is true, Eloise wins and otherwise Abelard does.

In this case Abelard has a winning strategy since he can pick $a = 1, b = 2$. Since $a < b$, picking that is out for Eloise. But if she picks a c, either $a \not< c$ or $c \not< b$ and Abelard wins.

Definition: Let A be a *closed* first order formula made up from literals using $\vee, \wedge, \forall, \exists$. M is an interpretation of the language of A. We define the game $G^M_A = G_A$ as follows:

1) If A is a literal, then the game is over and Eloise wins iff A is true in M.

2) If $A = B \wedge C$ then Abelard picks one of the games G_B, G_C. which is then played.

3) If $A = B \vee C$ then Eloise picks one of the games G_B, G_C.

4) If $A = (\exists x)B(x)$ then Eloise picks an element $a \in \|M\|$ and the game $G_{B(a)}$ is then played.

5) If $A = (\forall x)B(x)$ then Abelard picks an element $a \in \|M\|$ and the game $G_{B(a)}$ is then played.

Theorem: $M \models A$ iff Eloise has a winning strategy for G_A.

Proof: This is clear if A is a literal.

1) Suppose A is $B \wedge C$. Then $M \models A$ iff $M \models B$ and $M \models C$ iff (induction hypothesis) Eloise has winning strategies for both G_A and G_B iff she has a winning strategy for G_A. (She needs both winning strategies as it is Abelard who chooses which game to play).

2) Suppose A is $(\exists x)(B(x))$. If $M \models A$ then there is an element $a \in \|M\|$ such that $M \models B(a)$, but then Eloise has a winning strategy for $G_{B(a)}$ and hence for G_A. The converse is similar.

The cases for $\vee, \forall$ are similar. $\square$

Games of Partial Information

Chess is a game of full information. At each stage, each player knows the full situation. By contrast, Bridge is a game of partial information. At the start a player knows only her own hand. Logics corresponding to games of partial information have been investigated by a number of researchers including Hintikka, Sandu, Hodges, etc. Such logics go back to the so called Henkin quantifiers, but the current name for them is IF-logic, or Independence Friendly Logic.

Consider the formula $A = (\forall x)(\forall y)(\exists z/y)R(x, y, z)$. We define its meaning via a game. In this game, Abelard again picks elements a, b for x, y, but then Eloise has to choose c for z without knowing the value of b, and still make sure that $R(a, b, c)$ holds.

Now consider $B = (\forall x)(\forall y)(\exists z/y)(\exists u/x)((x = y \leftrightarrow z = u) \wedge (z \neq c \wedge u \neq c))$.

We claim that B holds in a structure M iff $\|M\|$ is infinite.

For if indeed $\|M\|$ is infinite, then there is a 1-1 function f from $\|M\|$ into $\|M\| - \{c\}$. Eloise can then pick $z = f(x)$ (not knowing y) and $u = f(y)$. This clearly is a winning strategy. Conversely, to win, she needs to have a 1-1 function f from $\|M\|$ into $\|M\| - \{c\}$.

It is known from the compactness theorem that no first order formula holds *exactly* in infinite models. Hence B above cannot be expressed in first order logic.

However, we can also consider *Finite Information Logic* [11] in which we study games where Eloise only has a *finite* amount of information about Abelard's moves. In that case we get a decidable sub-logic of first order logic, with the finite model property.

Instead of talking about ignorance, we could speak about knowledge, and in $(\forall x)(\forall y)(\exists z/x)$, instead of saying that Eloise does *not* know x, we could just as easily say that Eloise's knowledge is restricted to the value of y. Instead of concentrating on what Eloise does *not* know we concentrate on what she *does*. Similar restrictions might of course apply to Abelard in case he too has a move which follows the move of Eloise. (However, since we shall define the semantics in terms of the winning strategy of Eloise, it will turn out that the restrictions on Abelard do not enter into the semantics. To be sure of winning, Eloise must allow the possibility that Abelard makes a good move by luck.)

Now we introduce an innovation which will turn out to be interesting. IF-logic allows Eloise to know the value of x, or of y or of both or neither. Could we consider other possibilities? E.g. suppose x, y are integers. We might restrict Eloise to know the value of their *sum*, but not x, y themselves. Or for another, real life example, suppose you meet on the airplane an attractive woman who tells you only her first name (until she knows you better). Now if x is the name variable whose value is Eloise Dzhugashvili and she only tells you 'Eloise', then you do not know x but neither are you completely ignorant of it. You know it *in part*.

This opens up the possibility of more general kinds of knowledge of the values of variables than allowed by IF-logic and we will see that it leads to interesting possibilities.

As usual we have variables, predicate symbols, certain special function symbols. Atomic formulas are defined as usual. Literals are atomic formulas or their negations. For simplicity we will apply negation only to atoms.

Definition 1 1 *Literals are formulas of PI.*

2a *If $\varphi(\vec{x}, y)$ is a formula of PI and f is one of the special function symbols, then $(\exists y /\!\!/_{f(\vec{x})})\varphi(\vec{x}, y)$ is a formula of PI.*

2b *If $\varphi(\vec{x}, y)$ is a formula of PI and f is one of the special function symbols, then $(\forall y /\!\!/_{f(\vec{x})})\varphi(\vec{x}, y)$ is a formula of PI.*

3a *If $\varphi(\vec{x}), \theta(\vec{x})$ are formulas of PI then $\varphi(\vec{x}) \vee /\!\!/_{f(\vec{x})} \theta(\vec{x})$ is a formula of PI.*

3b *If $\varphi(\vec{x}), \theta(\vec{x})$ are formulas of PI then $\varphi(\vec{x}) \wedge /\!\!/_{f(\vec{x})} \theta(\vec{x})$ is a formula of PI.*

Intuitively, the $\exists y$ in $(\exists y /\!\!/_{f(\vec{x})})\varphi(\vec{x}, y)$ is Eloise's move but because of the restriction $/\!\!/_{f(\vec{x})}$ she only knows $f(\vec{x})$ when she makes her move. We may, more generally, allow her also to know the values of two or more functions f, g of $\vec{x}$ so that in the extreme case she could know all the projection functions and hence know $\vec{x}$ precisely. That case corresponds to our usual first order logic. In an intermediate case, she could know *some* of the projection functions on $\vec{x}$, i.e. some but not all of the variables in $\vec{x}$. That case corresponds to IF-logic.

In $(\forall y /\!\!/_{f(\vec{x})})\varphi(\vec{x}, y)$ the move is Abelard's and he too is restricted in a similar way. (However, this restriction will not affect the semantics which depends only on the winning strategies of Eloise.)

Let us consider $\varphi(\vec{x}) \vee /\!\!/_{f(\vec{x})} \theta(\vec{x})$. Since we have a disjunction here, it is for Eloise to choose which of the two formulas φ, θ to play. But when she chooses, she only knows the value $f(\vec{x})$ or perhaps more than one such value, but her knowledge of $\vec{x}$ might not be complete.

On the other hand, in $\varphi(\vec{x}) \wedge /\!\!/_{f(\vec{x})} \theta(\vec{x})$ the move is Abelard's but the restrictions are similar to those in 3a above.

Compositional semantics can be defined for PI in just the same way as they have been defined for IF-logic by Hodges [6], Väänänen [12], etc. Moreover PI-logic can be interpreted into second order logic in the same way. [7, 8, 9, 10] give a game theoretic interpretation of various logics, including classical and intuitionistic (but not IF-logic).

Now we come to a special kind of PI-logic where the special functions f allow only a *finite* amount of information about the arguments. Thus if a, b are integers and Eloise has to make a choice based on them, she might be allowed only to know whether $a < b$ or whether $a + b$ is odd, or whatever. Knowing the precise value of a, b or even of $a + b$ is out of the question.

Why consider such a restricted case? We have two reasons. One is that this special case of PI-logic which we shall call FI-logic, or *finite information logic* has very elegant logical properties. The other is that since quantifiers correspond to moves in games, the games which FI-logic represents arise all the time in social algorithms and are deeply related to how social human interactions work.

For example a passport official at an airport only wants to know whether you have a valid visa or not. If you do, she lets you in, if not, she sends you back on the next flight. Or perhaps she classifies you among four classes, those who are citizens, those who come from friendly countries whose citizens do not require a visa, those who have a visa, and the remaining who are the ones sent back. In any case she only wants a finite amount of information about the variable, namely you.

Or a young man looking for a date might want to know if the prospective date is blonde or brunette. If she is blonde, he is not interested, he wants to date brunettes only. If she does have dark hair, he wants to know if she is tall. If not, he is again not interested. So he seeks a finite amount of information about the prospective date. Naturally she may have similar questions about him. But each will seek only a finite amount of information.

We repeat the definitions which we had above for formulas of PI-logic, indicating where the difference arises between PI-logic in general and its special case, FI. Since only a finite amount of information is available at each step, it could easily be represented by one or more booleans, i.e. by formulas. Thus our special functions f drop out. Our main result is Theorem 11 which says that every consistent FI-sentence has a finite model. We use a strong form of this result to show that FI is exactly the existential-universal fragment of first order logic, if considered as a classical logic. However, FI is actually a non-classical logic with a rich many-valued semantics (this aspect will not be pursued in this paper). The reduction to first order logic is non-trivial in the sense that there is a trade-off: every $\exists\,\forall$ expression can be exponentially longer than its FI representation.

Before we formally define the *finite information logic* FI in defnitions 3 and 4 and discuss its semantics it makes sense to pay attention to what kind of θ we allow in $/\!\!/_\theta$, as the following informal result demonstrates:

Lemma 2 *The following conditions are intuitively equivalent in any model $\mathfrak{A}$ with at least two elements, whatever sentence θ is:*

1. $\mathfrak{A} \models (\forall x)(\exists y /\!\!/_{(x=c\vee\theta)})(y \neq x)$.
2. $\mathfrak{A} \models \neg\theta$

Proof. Let $\mathfrak{A}$ have two elements c, d (and perhaps others). Suppose θ is true. Then the information that $(x = c \vee \theta)$ is true tells $\exists$ nothing about x. Also the information that $(x = c \vee \theta)$ is false tells nothing because this information is impossible, i.e. never given in this case. Thus in this case y has to be chosen completely independently of x and $\exists$ cannot possibly have a winning strategy for choosing $y \neq x$.
On the other hand, suppose θ is false (and $\exists$ knows it). Then she can make the following inference: If I am told that $(x = c \vee \theta)$ is true, I know that it is true because $x = c$, and then I know what x is. So I can choose $y = d$. If I am told that $(x = c \vee \theta)$ is false, I know it is because $x \neq c$, and I can choose $y = c$. Thus $\exists$ can use the strategy, choose $y = d$ if θ is true and $y = c$ otherwise. □

In the proof we used the assumption that although the information that $\exists$ has is limited as to the values of the variables, $\exists$ can act as if she knows "generally known" things. For example, it follows that if $\exists$ has a winning strategy, she knows what it is. Also, if it is known that $\neg\theta$ (in a given model), then $\exists$ knows it too. But this is only metaphorical. Clearly if there is a strategy which works because certain sentences are true then $\exists$ can play it and does not need to "know" that these sentences are true.

Lemma 2 shows that if we allow θ in $/\!\!/_\theta$, we are committed to have also the negation of θ. On the other hand, games of imperfect information may very well be non-determined. Therefore we should be cautious with negation.

In social software it seems that the information we use in decisions is often atomic ("man", "woman") or existential ("has a ticket", "has a visa, which is valid") or boolean combinations of such ("is retired or has a serious defect in vision"). Accordingly we start by allowing θ in $/\!\!/_\theta$ to be any boolean combination of existential formulas.

Definition 3 *The set of formulas of FI is defined as follows:*

(1) *Atomic and negated formulas are FI-formulas. (Allowing function symbols does no harm as we indicate later.)*

(2) *If $\varphi(\vec{x})$ and $\psi(\vec{x})$ are FI- formulas and $\theta(\vec{x})$ is a boolean combination of existential formulas, then*

$$\varphi(\vec{x}) \wedge /\!\!/_{\theta(\vec{x})} \psi(\vec{x})$$

and

$$\varphi(\vec{x}) \vee /\!\!/_{\theta(\vec{x})} \psi(\vec{x})$$

are FI-formulas.

(3) *If $\varphi(\vec{x}, y)$ is an FI-formula and $\theta(\vec{x})$ is a boolean combination of existential formulas, then*

$$(\forall y /\!\!/_{\theta(\vec{x})}) \varphi(\vec{x}, y)$$

and

$$(\exists y /\!\!/_{\theta(\vec{x})}) \varphi(\vec{x}, y)$$

are FI-formulas.

For simplicity we shall leave out function symbols, but as we shall indicate later, the presence of functions need not interfere with some of our main results.

We have already given an intuitive explanation of FI. We now define a more formal semantics for FI. Suppose $\mathfrak{A}$ is a model and X is a set of functions s such that

(1) $\mathrm{dom}(s)$ is a finite set of variables

(2) $s, s' \in X \Longrightarrow \mathrm{dom}(s) = \mathrm{dom}(s')$

(3) $\mathrm{ran}(s) \subseteq A$.

Intuitively X is a set of plays i.e. assignments of values to variables. To incorporate partial information we have to consider sets of plays rather than mere individual plays. A partition $X = X_0 \cup X_1$ is $\theta(\vec{x})$-**homogeneous**, where $\theta(\vec{x})$ is first-order, if for all $s, s' \in X$

$$(\mathfrak{A} \models_s \theta(\vec{x}) \Longleftrightarrow \mathfrak{A} \models_{s'} \theta(\vec{x})) \Longrightarrow (s \in X_0 \Longleftrightarrow s' \in X_0).$$

Let

$$\begin{aligned} X[a:y] &= \{(s\backslash\{\langle y,b\rangle : b\in A\})\cup\{\langle y,a\rangle\} : s\in X\} \\ X[A:y] &= \{s\cup\{\langle y,a\rangle\} : s\in X, a\in A\}. \end{aligned}$$

We define the concept

$$\mathfrak{A}\models_X \varphi$$

for $\varphi \in FI$ as follows:

Definition 4 **(S1)** *$\mathfrak{A}\models_X \varphi$ iff $\mathfrak{A}\models_s \varphi$ for all $s\in X$, if φ is atomic or negated atomic.*

(S2) *$\mathfrak{A}\models_X \varphi(\vec{x})\wedge\!/\!\!/_{\theta(\vec{x})}\psi(\vec{x})$ iff $\mathfrak{A}\models_X \varphi(\vec{x})$ and $\mathfrak{A}\models_X \psi(\vec{x})$. ($\theta(\vec{x})$ plays no role)*

(S3) *$\mathfrak{A}\models_X \varphi(\vec{x})\vee\!/\!\!/_{\theta(\vec{x})}\psi(\vec{x})$ iff there is a $\theta(\vec{x})$-homogeneous partition $X = X_0\cup X_1$, such that $\mathfrak{A}\models_{X_0}\varphi(\vec{x})$ and $\mathfrak{A}\models_{X_1}\psi(\vec{x})$.*

(S4) *$\mathfrak{A}\models_X (\exists y/\!\!/_{\theta(\vec{x})})\varphi(\vec{x},y)$ iff there is a $\theta(\vec{x})$-homogeneous partition $X = X_0\cup X_1$, and a_0, a_2 such that $\mathfrak{A}\models_{X_0[a_0:y]}\varphi(\vec{x},y)$ and $\mathfrak{A}\models_{X_1[a_1:y]}\varphi(\vec{x},y)$.*

(S5) *$\mathfrak{A}\models_X (\forall y/\!\!/_{\theta(\vec{x})}))\varphi(\vec{x},y)$ iff*

$$\mathfrak{A}\models_{X[A:y]}\varphi(\vec{x},y)$$

($\theta(\vec{x})$ plays no role).

There is an asymmetry between $\wedge\!/\!\!/_{\theta(\vec{x})}$ and $\vee\!/\!\!/_{\theta(\vec{x})}$ on one hand and between $(\forall y/\!\!/_{\theta(\vec{x})})$ and $(\exists y/\!\!/_{\theta(\vec{x})})$ on the other hand. This is because in this paper we consider truth from the point of view of $\exists$ only, i.e. "classically". Thus we are concerned about the knowledge that $\exists$ has. As $\exists$ has to be prepared to play against all strategies of $\forall$, $\exists$ has to consider also the case that $\forall$ plays "accidentally" with perfect information. If we considered FI "non-classically" the symmetry would be preserved.

Lemma 5 *Suppose $\mathfrak{A}\models_{\{\emptyset\}}\varphi$. Then $\exists$ has a winning strategy in the obvious semantic game, namely, while $\exists$ plays, she keeps $\mathfrak{A}\models_X\varphi$ and "the play is $\in X$" remains true.*

Proof.

(G1) Suppose we are at an atomic or negated atomic formula φ. Since $\mathfrak{A}\models_X\varphi$ and the play is in X, $\exists$ wins by (S1).

(G2) We are at $\varphi(\vec{x})\wedge\!/\!\!/_{\theta(\vec{x})}\psi(\vec{x})$. Now $\forall$ plays choosing, say, $\varphi(\vec{x})$. We use (S2) to conclude $\mathfrak{A}\models_X\varphi(\vec{x})$.

(G3) We are at $\varphi(\vec{x}) \vee\!\!/\!\!/_{\theta(\vec{x})} \psi(\vec{x})$. We can by (S3) divide $X = X_0 \cup X_1$ in a $\theta(\vec{x})$-homogeneous way and $\mathfrak{A} \models_{X_0} \varphi(\vec{x})$ and $\mathfrak{A} \models_{X_1} \psi(\vec{x})$. The play is in X so it is in one of X_0 and X_1, but $\exists$ does not know in which. We let $\exists$ make the choice on the basis of the following inference. If $\theta(\vec{x})$ is true and some $\vec{x}\,'$ in X_0 satisfies $\theta(\vec{x}\,')$, then she chooses X_0. In this case homogeneity gives that the play is in X_0 and we also have $\mathfrak{A} \models_{X_0} \varphi(\vec{x})$. If $\theta(\vec{x})$ is true and some $\vec{x}\,'$ in X_1 satisfies $\theta(\vec{x}\,')$, then she chooses X_1. Again homogeneity gives that the play is in $\in X_1$ and we also have $\mathfrak{A} \models_{X_1} \varphi(\vec{x})$. Similarly, if $\theta(\vec{x})$ is false and some $\vec{x}\,'$ in X_0 satisfies $\theta(\vec{x}\,')$, then she chooses X_1, otherwise X_0.

(G4) We are at $(\forall y /\!\!/_{\theta(\vec{x})}))\varphi(\vec{x}, y)$. $\exists$ knows $\mathfrak{A} \models_{X[A:y]} \varphi(\vec{x}, y)$ and the play so far is in X. Whatever $\forall$ plays, the play is in $X[A : y]$.

(G5) We are at $(\exists y /\!\!/_{\theta(\vec{x})})\varphi(\vec{x}, y)$. There is a $\theta(\vec{x})$-homogeneous partition $X = X_0 \cup X_1$, and a_0, a_1 such that $\mathfrak{A} \models_{X_0[a_0:y]} \varphi(\vec{x}, y)$ and $\mathfrak{A} \models_{X_1[a_1:y]} \varphi(\vec{x}, y)$. As in the case of disjunction, player $\exists$ chooses a_0 or a_1 according to whether some $\vec{x}\,'$ in X_0 satisfies $\theta(\vec{x}\,')$ or not.

□

Examples 6 1° $(\forall x /\!\!/)(\exists y /\!\!/_{P(x)})(x = y)$ *says that both* P *and its complement have at most one element*

2° $(\forall x /\!\!/)(\exists y /\!\!/_{P(x)})(x \neq y)$ *says that both* P *and its complement are non-empty.*

For instance, in 2°, if both P and its complement are non-empty, then knowing the truth value of $P(x)$, $\exists$ can choose a y such that $P(y)$ has the opposite truth value. But this strategy (or any other) will not be available to her if either P or its complement is empty.

Lemma 7 *If* $\mathfrak{A} \models_X \varphi$ *and* $X_0 \subseteq X$, *then* $\mathfrak{A} \models_{X_0} \varphi$.

Proof. Immediate from the definition by induction on ϕ. □

Lemma 8 *Every* FI*-sentence is equivalent to a first order sentence (i.e. holds in the same models).*

Proof. Suppose $\phi \in FI$. Let n be the length of ϕ. It suffices to show that if $M \models \phi$ and player II has a winning strategy in the n-move Ehrenfeucht-Fraisse game [1, 4] on M and M', then $M' \models \phi$. Suppose X is a set of interpretations of a set V of variables in M and

$\Pi = \{\pi_s : s \in X\}$ is a set of partial isomorphisms $M \to M'$ such that $ran(s) = dom(\pi_s)$ for $s \in X$. Then $\{\pi_s \circ s : s \in X\}$ is a set of interpretations in M' of the variables in V. We write $X' = \Pi \circ X$. Let us assume that player II has a winning strategy in the Ehrenfeucht-Fraisse game on M and M' of n rounds. The different plays of this game, player II following her winning strategy, form in a natural way a tree T. If t is a node of height t of the tree T, that is, a play of i rounds, we denote by s_t the interpretation this play gives in M to the variables $x_1, ..., x_i$. Respectively, s'_t is the corresponding interpretation given in M to the variables $x_1, ..., x_i$. Since player II is playing a winning strategy, the mapping $s_t(x_j) \mapsto s'_t(x_j)$ is a partial isomorphism $M \to M'$. It T' is a subtree of T, let $X_i(T')$ be a set of all s_t for t in T' of height i. Let $X'_i(T')$ be the corresponding set of s'_t.

Claim. For all subtrees T' of T: $M \models_{X_i(T')} \phi(x_1, ..., x_i) \iff M' \models_{X'_i(T')} \phi(x_1, ..., x_i)$ for $\phi(x_1, ..., x_i)$ of quantifier rank $\leq n - i$.

The proof of the claim goes by induction on ϕ. The case of atomic formulas and conjunctions are trivial. For disjunction and existential quantifier it suffices to notice that if $X_i(T') = X_{i,0}(T') \cup X_{i,1}(T')$ is a $\theta(\vec{x})$-homogeneous partition, then $X'_i(T') = X'_{i,0}(T') \cup X'_{i,1}(T')$ is likewise a $\theta(\vec{x})$-homogeneous partition. For this $\theta(\vec{x})$ need not be a Boolean combination of existential formulas as long as it is first order. For universal quantifier we let player I try all possible elements and apply the induction hypothesis to the resulting new subtree of T. $\square$

Let $FI(FO)$ denote the extension of FI where any first-order θ is allowed to occur in $/\!\!/_\theta$. Lemma 2 implies that $FI(FO)$ contains all of first-order logic.

Corollary 9 *$FI(FO) = FO$.*

A first order formula is existential-universal $\exists\forall$ if it is of the form

$$(\exists x_1) \ldots (\exists x_n)(\forall y_1) \ldots (\forall y_m)\varphi$$

where φ is quantifier-free. A formula is Δ_1 if it is equivalent to an $\exists\forall$-formula and its negation is too. An example of a Δ_1 formula is

$$(\exists x_1)(\exists x_2)(x_1 \neq x_2) \wedge (\forall x_1)(\forall x_2)(\forall x_3)(x_1 = x_2 \vee x_1 = x_3 \vee x_2 = x_3).$$

which says that there exactly three elements. Boolean combinations of existential formulas (and of Δ_1 formulas) are, of course, Δ_1.

Lemma 10 *The following conditions are equivalent for any first order sentence φ:*

(1) *φ is equivalent to an $\exists\forall$-formula.*

(2) *If* $\mathfrak{A} \models \varphi$ *and* $\mathfrak{A}$ *is the union of a chain* $\mathfrak{A}_\alpha$ $(\alpha < \beta)$ *of models, then there is an* $\alpha < \beta$ *such that* $\mathfrak{A}_\alpha \models \varphi$.

(3) *If* $\mathfrak{A} \models \varphi$, *then there is* $C \subseteq A$ *finite such that for all* $\mathfrak{D}$ *with* $C \subseteq D \subseteq A$ *we have* $\mathfrak{D} \models \varphi$.

Proof. Clearly (1) $\rightarrow$ (3) $\rightarrow$ (2). We prove (2) $\rightarrow$ (1). By (2) the class of models of the sentence $\neg\varphi$ is closed under unions of chains of models. By the Łoś-Suszko lemma, $\neg\varphi$ is universal-existential, whence φ is equivalent to an $\exists\forall$ formula. $\square$

Theorem 11 *Every FI-sentence has the finite model property.*

Proof. We prove condition (3) of Lemma 10. We use induction on φ to prove:

($\star$) If $\mathfrak{A} \models_X \varphi$, where $X \subseteq {}^V A$, then there is a finite $A_0 = A_0(\mathfrak{A}, X, \varphi)$, s.t. for all $\mathfrak{A}_1 \subseteq \mathfrak{A}$ with $A_0 \subseteq A_1 \subseteq A$ we have $\mathfrak{A}_1 \models_{X \cap {}^V A_1} \varphi$.

(S1) φ is atomic or negated atomic. We can choose $\mathfrak{A}_0$ to be any non-empty subset of A.

(S2) Conjunction: We can let $A_0(\mathfrak{A}, X, \varphi \wedge \psi) = A_0(\mathfrak{A}, X, \varphi) \cup A_0(\mathfrak{A}, X, \psi)$ and this clearly works.

(S3) Disjunction: Suppose $\mathfrak{A} \models_X \varphi(\vec{x}) \vee\!\!/\!\!/_{\theta(\vec{x})} \psi(\vec{x})$. Let $X = X_0 \cup X_1$ such that $\mathfrak{A} \models_{X_0} \varphi(\vec{x})$, $\mathfrak{A} \models_{X_1} \psi(\vec{x})$ and the partition is $\theta(\vec{x})$-homogeneous. Remember that $\theta(\vec{x})$ is Δ_1. Let $\mathfrak{A}_1^*$ be finite such that $A_0(\mathfrak{A}, X, \varphi) \cup A_0(\mathfrak{A}, X, \psi) \subseteq \mathfrak{A}_1^* \subseteq \mathfrak{A}$ and $\mathfrak{A}_1^* \subseteq \mathfrak{A}_2 \subseteq \mathfrak{A}$ implies for all $s \in X$

$$\mathfrak{A}_2 \models_s \theta(\vec{x}) \iff \mathfrak{A} \models_s \theta(\vec{x}).$$

For such $\mathfrak{A}_2$ we have $\mathfrak{A}_2 \models_{X_0 \cap {}^V A_2} \varphi$ and $\mathfrak{A}_2 \models_{X_1 \cap {}^V A_2} \psi$. Moreover, the partition of $X \cap {}^V A_2$ to $X_0 \cap {}^V A_2$ and $X_1 \cap {}^V A_2$ is clearly $\theta(\vec{x})$-homogeneous.

(S4) Universal quantification: Suppose $\mathfrak{A} \models_X (\forall y)\varphi(\vec{x}, y)$. Thus $\mathfrak{A} \models_{X[A:y]} \varphi(\vec{x}, y)$. Choose $A_0 = A_0(\mathfrak{A}, X[A : y], \varphi((\vec{x}, y))$. If $\mathfrak{A}_1 \subseteq \mathfrak{A}$ with $A_0 \subseteq A_1 \subseteq A$, then $X[A : y] \cap {}^{(V \cup \{y\})} A_1 = (X \cap {}^V A_1)[A_1 : y]$, whence $\mathfrak{A}_1 \models_X (\forall y)\varphi(\vec{x}, y)$.

(S5) Existential quantification: $\mathfrak{A} \models_X (\exists y /\!\!/_{\theta(\vec{x})})\varphi(\vec{x}, y)$. Let $X = X_0 \cup X_1$ be $\theta(\vec{x})$-homogeneous and a_0, a_1 such that $\mathfrak{A} \models_{X_0[a_0:y]} \varphi(\vec{x}, y)$ and $\mathfrak{A} \models_{X_1[a_1:y]} \varphi(\vec{x}, y)$. Remember that $\theta(\vec{x})$ is Δ_1. Let $\mathfrak{A}_1^*$ be finite such that $A_0(\mathfrak{A}, X[a_0; y], \varphi(\vec{x}, y)) \cup A_0(\mathfrak{A}, X[a_1; y], \varphi(\vec{x}, y)) \cup \{a_0, a_1\} \subseteq A_1^* \subseteq A$ and $\mathfrak{A}_1^* \subseteq \mathfrak{A}_2 \subseteq \mathfrak{A}$ implies for all $s \in X$

$$\mathfrak{A}_2 \models_s \theta(\vec{x}) \iff \mathfrak{A} \models_s \theta(\vec{x}).$$

For such $\mathfrak{A}_2$ we have $\mathfrak{A}_2 \models_{X\cap^V A_2} (\exists y /\!\!/_{\theta(\vec{x})})\varphi(\vec{x}, y)$. Why? Because $X\cap^V A_2 = (X_0\cap^V A_2)\cup (X_1\cap^V A_2)$ is a $\theta(\vec{x})$-homogeneous partition in $\mathfrak{A}_2$ and $\mathfrak{A}_2 \models_{(X_0\cap^V A_2)[a_0:y]} \varphi(\vec{x}, y)$, $\mathfrak{A}_2 \models_{(X_1\cap^V A_2)[a_1:y]} \varphi(\vec{x}, y)$.

□

The above theorem has an alternative proof using the concept of a D-structure (see [7, 8, 9, 10], which are related to [2]).

Example 12 *The sentence*

$$(\forall x /\!\!/)(\exists y /\!\!/_{x=x})(x \leq y)$$

says that the linear order $\leq$ has a last element. It has no negation in FI as the negation does not have the finite model property.

The finite model property would be true even if we allowed any Δ_1 formula θ to occur in $/\!\!/_\theta$. However, allowing $\exists\forall$-formulas θ leads us to new avenues: Let $FI(\exists\forall)$ be this generalization.

Theorem 13 *$FI(\exists\forall)$ does not have the finite model property.*

Proof. Let φ be the sentence

$$(\forall x /\!\!/)(\exists y /\!\!/_{\psi(x)})(y \neq x)$$

where $\psi(x)$ is the $\exists\forall$-formula

$$x = 0 \vee (\exists u)(\forall v)(v \leq u).$$

The vocabulary consists of $\leq$ and the constant 0. Let φ' be the conjunction of φ and the universal (hence FI) axioms of linear order.

Claim 1 $\langle \omega, \leq, 0\rangle \models \varphi'$. The task of $\exists$ is to choose $y \neq x$ knowing only whether $\psi(x)$ is true or not. She argues as follows: If I am told $\psi(x)$ is true, I know it is because $x = 0$, so I choose $y = 1$. If, on the other hand, I am told that $\psi(x)$ is not true, I know $x \neq 0$, so I choose $y = 0$.

Claim 2 φ' has no finite models. Suppose $\mathfrak{A} = \langle A, \leq, 0\rangle$ were one. Now $\psi(x)$ is true independently of x. So $\exists$ has no way of choosing $y \neq x$ on the basis of whether $\psi(x)$ is true or not. More formally, suppose $\mathfrak{A} \models_X \varphi'$, where $X = \{\emptyset\}$. Then $\mathfrak{A} \models_{X[A:x]} (\exists y /\!\!/_{\psi(x)})(y \neq x)$. Let $X[A : x] = X_0 \cup X_1$ be a $\psi(x)$- homogeneous partition and $a_0, a_1 \in A$ such that $\mathfrak{A} \models_{X_0[a_0:y]} y \neq x$ and $\mathfrak{A} \models_{X_1[a_1:y]} y \neq x$. Since $\psi(x)$ is always true, $X_0 = \emptyset$ or $X_1 = \emptyset$. Say $X_1 = \emptyset$. Thus $\langle x, a_0\rangle \in X_0$, whence $\mathfrak{A} \models_{X_0[a_0:y]} y = x$, a contradiction. □

Let $FI(IF)$ denote the extension of FI where any θ from IF-logic is allowed to occur in $/\!\!/_\theta$. We know that non-well-foundedness can be expressed in the IF-logic. Lemma 2 implies that $FI(IF)$ can express also well-foundedness. Thus $FI(IF)$ is not included in IF-logic.

The FI as we have defined it turns out to be translatable into first-order logic:

Theorem 14 *Every FI-sentence is equivalent to an* $\exists\forall$*-sentence, and vice versa, every* $\exists\forall$*-sentence is equivalent to an FI-sentence.*

Proof. One direction (FI $\mapsto \exists\forall$) follows from Theorem 11 and condition (3) of lemma 10. For the converse implication it suffices to notice that following are equivalent:

$$\mathfrak{A} \models (\exists x_1)\dots(\exists x_n)(\forall y_1)\dots(\forall y_m)\varphi$$
$$\mathfrak{A} \models_{\{\emptyset\}} (\exists x_1/\!\!/)\cdots(\exists x_n/\!\!/)(\forall y_1/\!\!/)\dots(\forall y_m/\!\!/)\varphi',$$

where φ' is obtained from φ by replacing each disjunction $\theta(\vec{x}) \vee \psi(\vec{x})$ by $\theta(\vec{x}) \vee /\!\!/_{\theta(\vec{x}),\psi(\vec{x})}\psi(\vec{x})$. Note that ϕ is quantifier-free, so its subformulas can occur in connection with $/\!\!/$. □

Theorem 15 *FI has an exponential compression relative to first order* $\exists\forall$ *logic.*

Proof. Consider the structure $\mathfrak{A}$ whose domain consists of all binary numerals. The predicate $C(x,y)$ means that $y = x + 1 \ mod \ 2^n$. Of course $0 \leq y < 2^n$. The predicate $P_i(x)$ for $i \leq n$ means that the i-th digit of x from the right is 1. Consider the formula $\theta = (\forall x)(\exists y/\!\!/_{P_1(x),\dots,P_n(x))}C(x,y))$. The formula says that $\exists$ can choose y knowing only the truth values of $P_i(x) : i \leq n$. θ is true in $\mathfrak{A}$, and remains true if we only take integers $< 2^n$. But it is not true in any (non-empty) sub-structure of size $< 2^n$. Such a substructure will always contain an x such that *the* y such that $C(x,y)$ is missing. Thus any $\exists\forall$ formula which was equivalent to θ would have to have at least 2^n quantifiers. □

However, note that if we use full first order logic to express θ we do not need exponential growth. For the formula $\phi = (\forall x)(\exists y)(\forall z)([\bigwedge_{i\leq n} P_i(x) \leftrightarrow P_i(z)] \to C(z,y))$ *is* equivalent to θ. Intuitively, if $\forall$ is allowed to *change* his move (from x tz) *after* $\exists$ has played hers, but satisfying the same booleans, then she is in effect restricted to what she could have done had she known *only* the values of the booleans.

We now show that every model of a FI-formula has a finite submodel of at most exponential size.

Theorem 16 *Let $\mathfrak{A} \models \varphi$ where the logical complexity of φ is n. Then $\mathfrak{A}$ has a submodel $\mathfrak{B}$ of φ of size at most $n2^n$.*

Proof.: Assume that φ is written so that all negations apply only to atoms, so that φ is constructed from literals using $\exists, \forall, \vee, \wedge$ only. Eloise has a winning strategy for the game corresponding to φ. For each move $\exists y$ of Eloise, consider the moves $\forall x /\!\!/_{P(x)}$ in whose scope y lies. There are at most n of such predicates $P(x)$ and the value of y is determined by the truth values of these $P(x)$. (y may be determined also by previous moves y' of Eloise, but these are also determined by these booleans P and therefore by *all* booleans, whether y is in their scope or not.) So consider the set $\mathcal{V}$ of all boolean vectors governing *any* move of Eloise. The cardinality of $\mathcal{V}$ is at most 2^n. For each move $\exists y_i$ of Eloise, her strategy gives a function f_i from $\mathcal{V}$ into A, the domain of $\mathfrak{A}$. Since Eloise has at most n moves, there are at most n functions, and the range of all these functions gives us a subset of A of size at most $n2^n$. Let this subset be B.

Consider the modified game where Abelard is allowed to move in A but Eloise is restricted to move in B. Clearly Eloise is free to use her former winning strategy and wins. Consider now a further restriction where Abelard is also restricted to B. Surely this does not harm Eloise and she still wins. But that means that if $\mathfrak{B}$ is the submodel corresponding to B, its size is at most $n2^n$ and $\mathfrak{B} \models \varphi$. $\square$

References

[1] A. Ehrenfeucht, An application of games to the completeness problem for formalized theories *Fundamenta Mathematicae*, 49 (1961) pp. 129-141.

[2] A. Ehrenfeucht, J. Geiser, C. E. Gordon and D. H. J. de Jongh. A semantics for non iterated local observation. Preprint 1971

[3] Andrzej Ehrenfeucht "Logic without Iterations" *Proceedings of the Tarski Symposium* (1974) pp. 265-268.

[4] R. Fraisse Sur quelques classifications des systemes de relations *Publ. Sci. Univ. Alger.* 1:35-182 (1954).

[5] Jaakko Hintikka and Gabriel Sandu. Informational independence as a semantical phenomenon. In *Logic, methodology and philosophy of science, VIII (Moscow, 1987)*, pages 571–589. North-Holland, Amsterdam, 1989.

[6] Wilfrid Hodges. Compositional semantics for a language of imperfect information. *Log. J. IGPL*, 5(4):539–563 (electronic), 1997.

[7] Rohit Parikh, "D-Structures and their Semantics", *Notices of the AMS* 19 (1972) A329.

[8] Rohit Parikh and John Mayberry, "D-structures and *-structures", *Notices of the AMS* 19 (1972) A454.

[9] D.H.J.de Jongh, N. Goodman and R. Parikh, "On Regular *-structures with Classical Theories", *J. Symb. Logic* 37 (1972) 777.

[10] Rohit Parikh, "D-Structures and their semantics", appeared in a volume dedicated to Johan van Benthem, University of Amsterdam, August 99.

[11] Rohit Parikh and Jouko Vaananen, "Finite information logic," *Annals of Pure and Applied Logic*, **134** (2005) 83-93.

[12] Jouko Väänänen, On the semantics of informational independence, Journal of the IGPL, 2002.