3. Plausible inference in an ordinal setting

The path between classical and probabilistic Bayesian reasoning

Contents

- Confidence preorderings
- Qualitative possibility theory
- Possibilistic logic
- Non-mononotonic reasoning with ordinal beliefs

ORDINAL UNCERTAINTY : CONFIDENCE PREORDERINGS

- In the set of propositions $\mathcal{E}$ define relation $\geq_L$

  A $\geq_L$ B means "A *at least as likely as* B"

- **Natural Properties**
  - non triviality : S $>_L$ Ø
  - reflexivity : A $\geq_L$ A
  - totality : A $\geq_L$ B or B $\geq_L$ A
  - transitivity: if A $\geq_L$ B and B $\geq_L$ C then A $\geq_L$ C
  - limit conditions S $\geq_L$ A $\geq_L$ Ø
  - monotony: **if** A $\subseteq$ C and D $\subseteq$ B **then** A $\geq_L$ B **implies** C $\geq_L$ D.

EXTRACTING BELIEFS FROM A CONFIDENCE RELATION

- Given a confidence preordering of events, the set of beliefs induced by this relation is

$$\mathcal{A}\,(\geq_L) = \{A : A >_L A^c\}$$

- If the confidence relation represents generic knowledge, then given a set of observations representing evidence on the current state of facts and modelled by an event C such that $C >_L \emptyset$, the set of beliefs induced by $(\geq_L)$ in context C is :

$$\mathcal{A}^E(\geq_C) = \{A : A \cap C >_L A^c \cap C\}$$

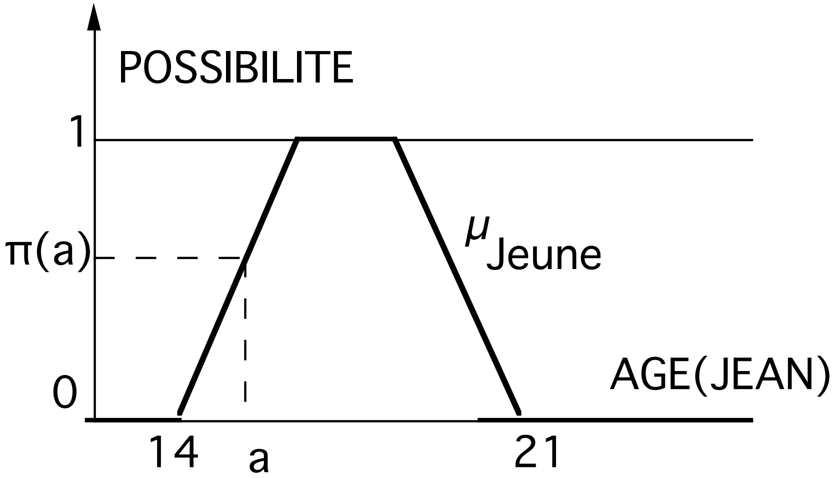
A SIMPLE ORDINAL REPRESENTATION OF INCOMPLETE KNOWLEDGE : **PLAUSIBILITY RANKINGS**

- **Idea** : refine a piece of information « $x \in E$ » by providing a ranking of states in S in terms of plausibility.
- **Definition** : Equip S with a complete preordering of states $\geq_\pi$:$s_1 \geq_\pi s_2$ means s_1 is more plausible, more normal, less surprising than s_2
- **Equivalent representation** : a well-ordered partition (E_1, $E_{2\ldots}E_n$) of S, where E_1 contains the most likely states, E_n the least likely ones
- A more expressive framework than disjunctive sets

ABSOLUTE QUALITATIVE REPRESENTATIONS OF PLAUSIBILITY RANKINGS

- a possibility distribution π maps the well-ordered partition $(E_1, E_{2 \ldots} E_n)$ to a plausibility scale L using L is a chain with top 1 and bottom 0.
- *A possibility distribution* π_x *is the representation of a state of knowledge: what an agent knows of the state of affairs x is*.
- **Conventions**
  - $\pi_x(s) = 0 \Leftrightarrow x = s$ is impossible, totally excluded *(not expressible with* $\geq_\pi$*)*
  - $\pi_x(s) = 1 \Leftrightarrow x = s$ is expected, normal, fully plausible, unsurprising
  - $\pi_x(s) > \pi_x(s')$ $\Leftrightarrow x = s$ more plausible than $x = s'$

POSSIBILITY DISTRIBUTIONS FROM LINGUISTIC INFORMATION



- Information item « JEAN IS YOUNG »
  - $\mu_{\mathbf{YOUNG}}$(a) = possibility (AGE(JEAN) = a)
  - YOUNG = FUZZY SUBSET OF POSSIBLE VALUES OF THE AGE OF JEAN

SUBJECTIVE ORDINAL UNCERTAINTY :

- Example : AGE OF THE PRESIDENT
- **partial ignorance :** $70 \leq x \leq 80$ (sets**,** intervals)
  - uniform possibility distributions
  - $\pi(x) = 1 \quad x \in [70,80]$
  - $= 0 \quad$ otherwise
- **partial ignorance with preferences**
  - he was *about* 10 when war started, hence
  - $72 >_\pi 71 \sim_\pi 73 >_\pi 70 \sim_\pi 74 >_\pi 75 >_\pi 76 >_\pi 77\ldots$
- *Note that this is uncertain evidence, not generic information…*
  - **But a plausibility ranking may also express generic information** : Flying birds $>_\pi$ Non-flying birds

Comparing information states

- **π' more specific than π in the wide sense**

  if and only if $\pi' \leq \pi$

In other words: any possible value in information state π' is at least as possible in information state π, *that is, π' is more informative than π*

- COMPLETE KNOWLEDGE : The most specific ones
  - $\pi(s_0) = 1$; $\pi(s) = 0$ otherwise
- IGNORANCE : $\pi(s) = 1, \forall s \in S$

- **Principle of minimal specificity** : if a state is not proved impossible it is possible: select the least informative epistemic state by *maximizing possibility degrees*.

MEASURES OF CONFIDENCE INDUCED BY POSSIBILITY DISTRIBUTIONS

- How confident are we that $x \in A \subset \mathbf{S}$?
- **The level of plausibility (possibility) *that $x \in A$***

$$\Pi(A) = \sup_{s \in A} \pi(s)$$

= to what extent **at least one** element in A is consistent with π (= possible)

- **The degree of certainty (necessity) that $x \in A$**

$$N(A) = \nu(\Pi(A^c)) \text{ (for instance } 1 - \Pi(A^c) \text{)}$$

= to what extent no element outside A is possible
= to what extent π implies A

- (ν *denotes an order-reversing map on L)*

COMPATIBILITY BETWEEN ORDINAL AND NUMERICAL SETTINGS

- *"A confidence measure g represents a confidence relation"* means $A \geq_L B$ iff $g(A) \geq g(B)$
- **Known examples :**
  - Probability measures induce comparative probabilities (not conversely) (Dubois, 1986)

    if $(A \cup B) \cap C = \emptyset$ **then** $A \geq_P B \Leftrightarrow A \cup C \geq_P B \cup C$
  - Possibility measures represent possibility relations

    for all A,B, $A \geq_\Pi B$ **implies** $A \cup C \geq_\Pi B \cup C$
  - Necessity measures represent epistemic entrenchment relations of Gärdenfors (Dubois, 1986, 1991)

    for all A,B, $A \geq_N B$ **implies** $A \cap C \geq_N B \cap C$

HISTORY and TERMINOLOGY of Possibility theories

- *Ordinal*
  - **Possibility relations** : David Lewis (1973) (modelling counterfactual information), Dubois(1986), Adam Grove (1988)
  - **Qualitative possibility distributions or plausibility rankings =** systems of spheres of Lewis and Grove.
  - **Necessity relations** : Dubois (1986); epistemic entrenchment relations in the field of belief revision of Gärdenfors (Dubois and Prade, 1991).
  - **Lattice-valued** possibility distributions (De Cooman)

HISTORY and TERMINOLOGY of Possibility theories

- *Numerical*
  - **Numerical impossibility measures :** Shackle's degrees of surprise (1950) ($1-\Pi$)
  - More recently Zadeh's (1978) coined the word "possibility measure": linguistic information as **fuzzy (disjunctive) sets**
  - Spohn's (ordinal conditional) **kappa functions** (integer exponents of infinitesimal probabilities)
  - Shafer's **consonant belief functions**
  - Special cases of **probability bounds** (Dubois and Prade, 1992)

QUALITATIVE POSSIBILISTIC REASONING

- *The simplest theory of reasoning with ordinal uncertainty : all information is contained in a plausibility ranking of states.*
- **Plausibility of events described in terms of possibility degrees.**
  - $\Pi(A)$ evaluates how unsurprising event A is
- **ASSUMPTION** for computing $\Pi(A)$: the current situation is the most normal where A is true

  $\Pi(B) \geq \Pi(A)$ means " the most plausible situation where B occurs is at least as plausible as the most plausible situation where A occurs"

  - *Comparing propositions on the basis of their most normal models*

Necessity degrees as grading acceptance

- By default the state of affairs is in the set E_1 of most plausible states.
- **Proposition** : $N(A) > 0$ iff $E_1 \subseteq A$:
  - It means A is true in **all** the normal situations
  - $N(A) > 0$ means : *A is accepted as an expected truth*
  - $N(A) = N(\text{not } A) = 0$: complete ignorance about A
- N(A) evaluates how strongly A is entrenched, an accepted belief,
  - **N(A) = level of acceptance of A**
  - **Note that** $N(A) > 0$ iff $N(A) > N(A^c)$ iff $\boldsymbol{\Pi(A) > \Pi(A^c)}$

THE POSSIBILISTIC REPRESENTATION OF BELIEF

- The complete preordering of states $\geq_\pi$ shares the set of propositions into 3 subsets
  - **Accepted beliefs** $\mathcal{A}$: $N(A) > 0$, i.e.,, $\mathbf{\Pi(A) > \Pi(A^c)}$
  - **Rejected beliefs**: $\mathcal{R}$: $\Pi(A) < 1$, i.e., $\mathbf{\Pi(A) < \Pi(A^c)}$
  - **Ignored beliefs**: $\mathcal{U}$: A such that $\Pi(A) = \Pi(A^c) = 1$
- *Like in classical logic: $\mathcal{A}$ is deductively closed*
  - incompleteness is captured
- *Unlike classical logic*: $\mathcal{A}$ is ranked in terms of certainty, $\mathcal{R}$ in terms of impossibility

REASONING WITH PLAUSIBILITY ORDERINGS

The set of propositions accepted in epistemic state $\geq_\pi$ is: $\mathcal{A}(\geq_\pi)$ = $\{A : \Pi(A) > \Pi(A^c)\}$

- **It is closed under deduction:**
  - $N(A) > 0$ and $A \subseteq B$ imply $N(B) > 0$
  - $N(A) > 0$ and $N(B) > 0$ imply $N(A \cap B) > 0$
- THIS IS DIFFERENT FROM PROBABILISTIC REASONING
  - BASED ON AVERAGING
  - $P(A) > a > 0.5$ cannot be interpreted as acceptance
  - The set $\{A : P(A) > P(A^c)\}$ is not closed

Possibility theory is the theory of defeasible acceptance

- DEFEASIBILITY : *If C is learned to be true, then the normal situations become the most plausible ones in C, and the accepted beliefs are revised accordingly*
- Given a plausibility ordering $\geq_\pi$ *representing generic information, and a context defined by evidence C, The accepted beliefs in context C are :*

$$\mathcal{A}^{E}(\geq_\pi) = \{A : \Pi(A \cap E) > \Pi(A^c \cap E)\}$$

  - ***This set is deductively closed for any non-impossible context ($\Pi(C) > 0$)***
- ***Theorem :*** *Given a confidence ordering $\geq_L$ then the set of beliefs induced by $\geq_L$ is deductively closed if and only if $\geq_L$ is a possibility ordering.*

Possibilistic logic: syntax

- A possibilistic knowledge base is a totally preordered set of sentences
- $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \ldots \mathcal{B}_m$ where $\mathcal{B}_i = \{(p_{ij}\ \alpha_i), j = 1, \ldots\}$ is the α_i-layer, priorities $\alpha_1 > \alpha_2 > \ldots \alpha_m$ lying in some ordinal scale.
- *Inference is a straightforward extension of classical inference :* $\mathcal{B} \vdash (p\ \alpha_{\mathbf{i}})$ iff there is an index **i** such that

1. $\{ p_{ij} \in \mathcal{B}_1 \cup \mathcal{B}_2 \ldots \cup \mathcal{B}_{\mathbf{i}} \}$ classically implies p
2. This set is not inconsistent.
3. $\{ p_{ij} \in \mathcal{B}_1 \cup \mathcal{B}_2 \ldots \cup \mathcal{B}_j \}$ does not imply p for $j < i$.

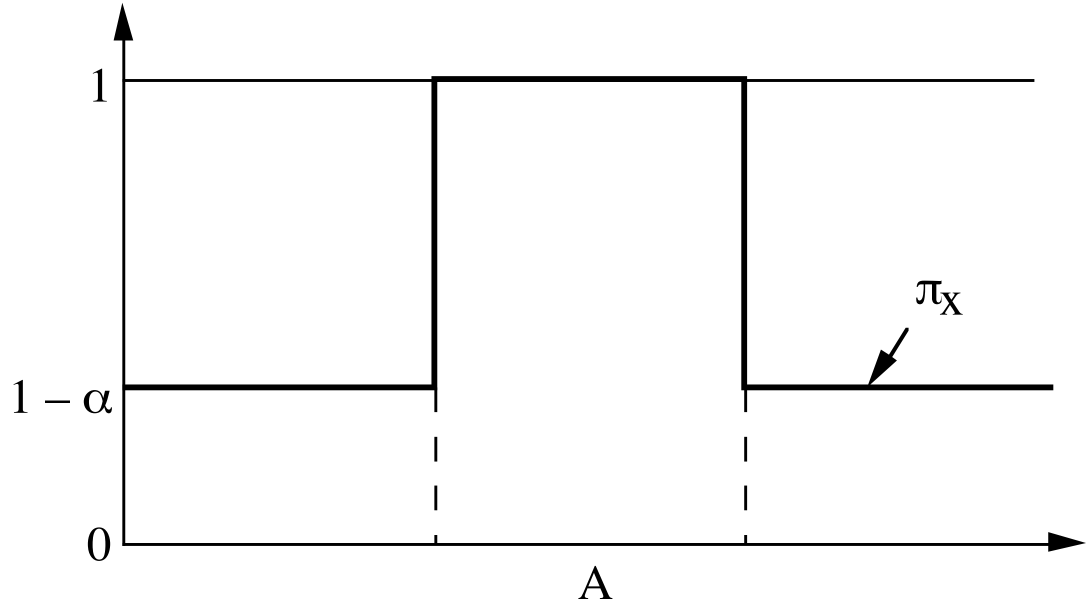
Possibilistic logic: proof method

- Basic principles
  - *The weight of a chain of inference is the weight of the weakest link*
  - The weight of the conclusion is the *weight of the strongest chain of inference that produces it*
- *Valid inference patterns*
  - *Modus ponens:* $\{(p, \alpha), (\neg p \vee q, \beta)\} \vdash (q, min(\alpha, \beta))$
  - *Resolution:* $\{(p \vee q, \alpha), (\neg p \vee r, \beta)\} \vdash (q \vee r, min(\alpha, \beta))$
  - *Fusion* $\{(p, \alpha), (p, \beta)\} \vdash (p, max(\alpha, \beta))$
- *Certainty of a conclusion p:* $max\{\alpha, \mathcal{B} \vdash (p, \alpha)\}$

Certainty qualification

- Attaching a degree of certainty to a proposition p:
- Denote « p is α-certain » by (p, α)
- It means $N(A) \geq \alpha$ where $A = [p]$ the set of models of p :
  - $N(A) \geq \alpha \Leftrightarrow \Pi(A^c) \leq 1 - \alpha \Leftrightarrow \pi(s) \leq 1 - \alpha, \forall s \notin A$
- The least informative possibility distribution sanctioning $N(A) \geq \alpha$ is :
  - $\pi_{(p, \alpha)}(s) =$ 1 if $s \in A$
  - $1 - \alpha$ if $s \notin A$

Certainty-qualification



- $\pi_{(p,\ \alpha)}(s) = \max(\mu_A, 1 - \alpha)$:
  - If $\alpha = 1$: $[(p, 1)] = [p] = A$
  - If $\alpha = 0$: $[(p, 0)] = [T] = S$ (tautology)

Possibilistic logic: semantic

A set of sentences $\mathcal{B}$ with priorities models certainty-qualified assertions;

- Constraints $\{N(A_{ij}) \geq \alpha_i, i = 1, n\}$ where A_{ij} is the set of models of p_{ij}
- (p, α) means « x is A is α-certain » : $N(A) \geq \alpha$
- Models of (p, α) form a fuzzy set:
  - $\pi_{(p, \alpha)}(s) =$ 1 if s satisfies p , $1 - \alpha$ if s does not satisfy p
- $\mathcal{B}$ is interpreted by the least specific possibility distribution on the set of interpretations obeying the constraints :

$$\pi_{\mathcal{B}} = \min_{ij} \max(\mu_{A_{ij}}, 1 - \alpha_i)$$

SOUNDNESS AND COMPLETENESS

- Semantic inference: $\mathcal{B} \models (p, \alpha)$ means $\pi_{\mathcal{B}} \leq \pi_{(p, \alpha)}$

- **Main theorem** : Possibilistic logic is sound and complete w.r.t. this semantics :

$$\mathcal{B} \models (p, \alpha) \text{ iff } \mathcal{B} \vdash (p, \alpha),$$

Inconsistency-Tolerant inference

- Degree of inconsistency of a possibilistic belief base:

$$\mathrm{Inc}(\mathcal{B}) = \max\{\alpha, \mathcal{B} \mid\!\!- (\perp, \alpha)\}$$

  – For all p, $\mathcal{B} \mid\!\!- (p, \mathrm{Inc}(\mathcal{B}))$,

- Inconsistency-Tolerant inference:

$$\mathcal{B} \mid\!\!-_{\mathrm{Pref}}\ p \text{ if } \mathcal{B} \mid\!\!- (p, \alpha) \text{ with } \alpha > \mathrm{Inc}(\mathcal{B}).$$

- The set of non-trivial consequences of $\mathcal{B}$ are those of the largest set $\{p_{ij} \in \mathcal{B}_1 \cup \mathcal{B}_2 \ldots \mathcal{B}_{\mathbf{i}}\}$ that is not inconsistent ($\mathrm{Inc}(\mathcal{B}) = \alpha_{i+1}$).

ORDINAL CONDITIONING

- Conditional possibility measures $\Pi(\cdot \mid C)$ are induced by the least informative possibility distribution on $C \neq \emptyset$ such that :

$$\Pi(A \cap C) = \min(\Pi(A \mid C), \Pi(C))$$

- It yields
  - $\Pi(A \mid C) \quad = 1$ if $A \neq \emptyset$, $\Pi(C) = \Pi(A \cap C) > 0$
  - $= \Pi(A \cap C)$ if $\Pi(A \cap C) < \Pi(C)$
  - $N(A \mid C) = 1 - \Pi(A^c \mid C)$ ordinal conditional necessity measures
  - $\pi(.\mid C)$ is the restriction of π to C + normalization, moving to 1 the possibility of normal states in C
- When $\Pi(C) > 0$, $N(A \mid C) > 0$ iff $\Pi(A \cap C) > \Pi(A^c \cap C)$

POSSIBILITY AS EXTREME PROBABILITY

- SPOHN's ORDINAL CONDITIONAL (KAPPA) FUNCTIONS: $\kappa(A)$ = disbelief in A
  - The higher $\kappa(A)$, the less likely.
- **Basic properties** :
  - $\kappa(A \cup B) = \min(\kappa(A), \kappa(B)) \in \mathcal{N}$ (integers)
  - $\kappa(S) = 0$
  - $\kappa(A \mid B) = \kappa(A \cap B) - \kappa(A)$ (conditioning rule)

- <u>**Probabilistic interpretation** :</u> there is some infinitesimal ε such that $\kappa(A) = n \Leftrightarrow P(A) \approx \varepsilon^n$
- $P(A \cup B) \approx \varepsilon^{\kappa(A)} + \varepsilon^{\kappa(B)} \approx \varepsilon^{\min(\kappa(A), \kappa(B))}$

POSSIBILITY AS EXTREME PROBABILITY

- **Possibilistic interpretation of kappa functions:**
- Transformation method : $\Pi_\kappa(A) = 2^{-\kappa(A)}$
  - Function Π_κ is a rational-valued possibility measure on [0, 1] with $\Pi_\kappa(A) > 0, \forall A \neq \emptyset$, hence $\kappa(A^c) = -\operatorname{Log}_2(1 - N(A))$
  - Then, $\Pi_\kappa(A)$ represents an order of magnitude whereby $\Pi_\kappa(A) > \Pi_\kappa(B)$ indicates that B as plausibility negligible in front of A
- It yields the product conditioning rule for possibility

$$\Pi_\kappa(A \mid B) = \quad \Pi_\kappa(A \cap B)/\Pi_\kappa(B)$$

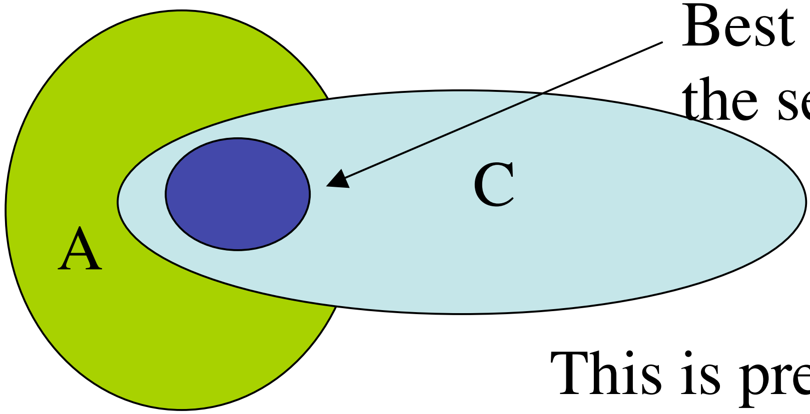
(special case of Dempster rule for belief functions)

- **Ranked belief bases**: (p, n) means $\kappa(\neg p) \geq n$ (integer), which is an alternative encoding of (p, α), that is $N(p) \geq \alpha = 1- 2^{-\kappa(\neg p)}$;

PLAUSIBLE CONSEQUENCEHOOD

Definition : A is a plausible consequence of evidence C, in the epistemic state π :

$$C \models_{\pi} A, \text{ iff } \Pi(A \cap C) > \Pi(A^c \cap C)$$



Best C-worlds in the sense of π

This is preferential inference a la Shoham

– **Theorem**: $C \models_{\pi} A$ iff A is true in the most plausible worlds where C is true:

$\forall s \in A$, if $\pi(s) = \Pi(A)$, then $s \in B$

Properties of plausible consequence $\models_\pi$

- A $\models_\pi$ A if A $\neq$ Ø (restricted reflexivity)
- Ø $\models_\pi$ A never holds (*From contradiction*, *nothing normally follows*)
- if A $\neq$ Ø, then A $\models_\pi$ Ø never holds (*consistency preservation*)
- The set {A: C $\models_\pi$ A} is **deductively closed:**
  - If A $\subseteq$ B and C $\models_\pi$ A then C $\models_\pi$ B (right weakening rule RW)
  - If A $\models_\pi$ B ; A $\models_\pi$ C then A $\models_\pi$ B $\cap$ C (Right AND)

Properties of plausible consequence $\models_\pi$

- If A $\models_\pi$ B and A $\cap$ B $\models_\pi$ C then A $\models_\pi$ C (*cut*, *weak transitivity*)
  - *(But if A normally implies B which normally implies C, then A may not imply C)*
- If A $\models_\pi$ B and if A$\models_\pi$ $\neg$C does not hold, then A $\cap$ C $\models_\pi$ B (***rational monotony RM***)
  - *If B is normally expected when A holds, then B is expected to hold when both A and C hold, unless it is that A normally implies not C.)*
- This is stronger than cautious monotony
  If A $\models_\pi$ B and A$\models_\pi$ C then A $\cap$ C $\models_\pi$ B
  - *(If B and C are normally expected when A holds, B is expected to hold when both A and C hold.)*

More properties

- If $A \cap B \models_{\pi} C$ then $B \models_{\pi} \neg A \cup C$ (*half of deduction theorem*)

- If $A \models_{\pi} C$; $B \models_{\pi} C$ then $A \cup B \models_{\pi} C$ (*Left OR*)

Plausible consequence in possibilistic logic

- Any possibility distribution π on a set of interpretation of a Boolean language can be represented by a possibilistic belief base $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \ldots \mathcal{B}_m$ such that $\pi = \pi_B$.
- The accepted beliefs in context where p is true are non-trivial consequences of $=$ B $\cup$ {(p, 1)}
- If $\pi = \pi_B$ C = [p], A = [q], then

$$C \models_{\pi} A \text{ if and only if } \{(p, 1)\} \cup \mathcal{B} \vdash_{Pref} q$$

Conservative plausible inference :

- Given an epistemic state π, consider the set $\mathbf{R}(\pi)$ of all epistemic states ρ more informative than π:
- DEFINITION: $C \models_{\forall} A$ iff **for all** $\rho \in \mathbf{R}(\pi)$ $\Pi_\rho(A \cap C) > \Pi_\rho(A^c \cap C)$
- We can restrict $\mathbf{R}(\pi)$ to all linear plausibility rankings that refine $\geq_\pi$ (Benferhat Dubois, Prade, 1999)
- This inference is weaker than $\models_\pi$
- It satisfies all properties of but rational monotony (only cautious monotony)
- It has the same properties as system P (if $C \neq \emptyset$).

REPRESENTATION THEOREM FOR POSSIBILISTIC ENTAILMENT

- Let $|\!\sim$ be a (consequence) relation on $2^S \times 2^S$
- It induces a relation $A > B$ iff $A \cup B \mid\!\sim B^c$ on subsets of S
  - **Lemma :** given a possibility distribution π, $A >_\Pi B$ iff $A \cup B \models_\pi B^c$
- **Theorem** (Benferhat et al, 1997): If the consequence relation $|\!\sim$ satisfies restricted reflexivity, consistency preservation, right weakening, **rational monotony,** Right AND and Left OR, then $A > B$ is the strict part of a possibilistic ordering of events
  - ***So a consequence relation satisfying the above properties is representable by possibilistic inference, and induces a complete plausibility ranking on the states.***

REPRESENTATION THEOREM FOR CONSERVATIVE POSSIBILISTIC ENTAILMENT

- Let $|\!\sim$ be a (consequence) relation on $2^S \times 2^S$
- It induces a relation $A > B$ iff $A \cup B \mid\!\sim B^c$ on subsets of S

- **Theorem** (Benferhat et al, 1999): If the consequence relation $|\!\sim$ satisfies restricted reflexivity, consistency preservation, right weakening, **cautious monotony,** Right AND and Left OR, then there exists a possibility distribution π such that $C \mid\!\sim A$ iff $C \models_{\forall} A$ for all linear refinements of π.

- So preferential inference of Kraus, Lehmann and Magidor has possibilistic semantics
- $C \mid\!\sim A$ can be interpreted as a three-valued conditional.

GENERIC CONDITIONAL KNOWLEDGE AS POSSIBILISTIC CONSTRAINTS

- A generic rule $p \rightarrow q$ can be modelled by a possibilistic constraint: $\Pi(p \wedge q) > \Pi(p \wedge \neg q)$
  - *If p is true then q is more plausible than ¬q*
  - *Examples are more likely than counterexamples*
- **This constraint delimits a set of possibility distributions on the set of interpretations of the language**
- **Example**
  - Birds fly: $b \rightarrow f$ = all ordinal possibility distributions such that $\Pi(b \wedge f) > \Pi(b \wedge \neg f)$
  - **Apply Minimal specificity principle**:
    - WOP= $\{E1 = [b \vee f], E2 = [b \wedge \neg f]\}$
    - $\Pi(b \wedge f) = \Pi(\neg b \wedge f) = \Pi(\neg b \wedge \neg f) = 1 > \Pi(b \wedge \neg f)$

POSSIBILISTIC SEMANTICS OF A SET OF GENERIC RULES

- **Set of default rules** $\Delta = \{p_i \to q_i, i = 1,n\}$
  - It defines a set of constraints on possibility distributions : $\Pi(p_i \wedge q_i) > \Pi(p_i \wedge \neg q_i), i = 1,..,n$
- **R**(Δ) = set of feasible π's with respect to Δ is the set of “models” of Δ.
  - Can write $\pi \models \Delta$ for “$p_i \models_\pi q_i$ for all rules in Δ”
  - ***Remark*** *: the models of a body of generic rules are epistemic states, not states of the world!!!*
- Compute the least informative possibility distribution π^* in **R**(Δ) such that $\pi^*(s)$ is maximal for each interpretation s

INFERENCE WITH A SET OF GENERIC RULES

- The *cautious* entailment of q based on evidence C = [p] and relying on knowledge Δ is modeled by

$$\Delta \models_{\forall} p \rightarrow q \quad \text{iff} \quad \forall \text{linear } \pi \models \Delta,\ p \models_{\pi} q$$

(i.e, $\Pi(p_i \wedge q_i) > \Pi(p_i \wedge \neg q_i)$, $i = 1,..,n$ imply $\Pi(p \wedge q) > \Pi(p \wedge \neg q)$, for all such linear distributions π)

- The *plausible* entailment of q based on evidence C = [p] and relying on knowledge Δ is modeled by

$$\Delta \models p \rightarrow q \text{ iff } p \models_{\pi *} q$$

(i.e, $\Pi^*(p \wedge q) > \Pi^*(p \wedge \neg q)$, for distribution π^*)

LINKS WITH OTHER APPROACHES

- **(Pearl) System Z**
  - **Toleration:** $p \to q$ is tolerated by Δ iff $\{p \wedge q, \neg p_i \vee q_i, i = 1,n\}$ is consistent

  = examples of $p \to q$ are not counterexamples of Δ

  - **Z-ranking** We can then partition Δ into $\Delta_1 \cup \Delta_2 \ldots \cup \Delta_k$ where $\forall i$, $\Delta_{i+1} \cup \Delta_{i+2} \cup \ldots \cup \Delta_k$ tolerate elements of Δ_i
- **Theorem** : The minimum specificity ranking of rules computed from Δ is exactly the Z-ranking proposed by Pearl, using toleration.

LINKS WITH OTHER APPROACHES

- **Lehmann's approach to non-monotonic reasoning with conditional assertions:**
  - The set $\{p \rightarrow q, \Delta \models A \rightarrow B\}$ of rules cautiously entailed by Δ under the possibility theory setting coincides with the **preferential closure** of Kraus Lehmann and Magidor

  - The set $\{p \rightarrow q, \Delta \models A \rightarrow B\}$ of rules plausibly entailed by Δ using π^* coincides with the **rational closure** of Lehmann

Encoding plausible reasoning in possibilistic logic

- Given a set of default rules $\Delta = \{p_i \rightarrow q_i, i = 1, n\}$ and the least informative possibility distribution π^* in $\mathbf{R}(\Delta)$, define the possibilistic belief base

$$\mathcal{B}_\Delta = \{(\neg p_i \vee q_i, \alpha_i), i = 1, n\}$$

where $\alpha_i = N^*(\neg p_i \vee q_i)$

Then $\Delta \models p \rightarrow q$ iff $\{(p, 1)\} \cup \mathcal{B}_\Delta \vdash_{Pref} q$

EXAMPLE

- IF BIRD THEN FLY
- IF PENGUIN THEN NOT-FLY
- IF PENGUIN THEN BIRD

- $\Delta = \{b \rightarrow f, p \rightarrow b, p \rightarrow \neg f\}$
- $\Pi(b \wedge f) > \Pi(b \wedge \neg f)$; $\Pi(p \wedge b) > \Pi(p \wedge \neg b)$; $\Pi(p \wedge \neg f) > \Pi(p \wedge f)$
- the min spec π^* is such that
    - abnormal situations: $f \wedge p$, $\neg b \wedge p$
    - less abnormal situations: $\neg f \wedge b$
    - normal situations: $\neg p \wedge b \wedge f$, $\neg p \wedge \neg b$
- **Ranking of rules:** $b \rightarrow f$ has less priority than others

$$1.\ \Pi(p\wedge b) > \Pi(p\wedge\neg b)\ ;$$
$$2.\ \Pi(b\wedge f) > \Pi(b\wedge\neg f)\ ;$$
$$3.\ \Pi(p\wedge\neg f) > \Pi(p\wedge f).$$

- *Step 1 : normal models* = unconstrained ones
  - $\neg(\,(p\wedge\neg b) \vee (b\wedge\neg f)\vee(p\wedge f)\,) = \mathbf{\neg p\wedge(\neg b \vee f)}$
    - =(Non-penguins that, either are birds or fly)
  - Since $(b\wedge f) \wedge\neg p\wedge(\neg b \vee f) = b\wedge f\wedge\neg p \neq \perp$, constraint 2 can be deleted.
- *Step 2 : sub-normal models* :
  - $\neg(\,p\wedge\neg b)\vee(p\wedge f)) \wedge (p\vee (b \wedge \neg f) = \mathbf{b \wedge \neg f}$
    - (Non-flying birds)
  - **Stop** : $b \wedge \neg f$ is consistent with $p\wedge b$ and $p\wedge\neg f$.
- *Abnormal Models* :
  - $\neg[(b \wedge \neg f) \vee (\neg p\wedge(\neg b \vee f)] = \mathbf{p\wedge(\neg b \vee f)}$
    - (penguins that either fly or are not birds)

Example

- **from the generic base of rules Δ one gets the possibilistic knowledge base**
- $K = \{(b \Rightarrow f, \alpha), (p \Rightarrow b\ \beta), (p \Rightarrow \neg f\ \beta)\}$,

where $\Rightarrow$ is a material implication and $\alpha < \beta$

- $K \cup \{b\} \vdash f$
- $K \cup \{p, b\} \vdash \neg f$
- $K \cup \{r, b\} \vdash f$, where r = red (not present in Δ)
  - This is due to rational monotony.
- *This behavior cannot be achieved in classical logic nor System P*

Conclusion

- Qualitative possibility theory is the natural framework for (non-monotonic) reasoning based on total plausibility orders of interpretations of a language.
- Possibilistic reasoning is a qualitative counterpart of Bayesian reasoning and Spohn kappa functions
- It relies on a natural notion of three-valued conditional due to De Finetti, in agreement with conditional probability, that can be used in all uncertainty theories.