Possibilistic Logic as Interpretability Logic

Hájek, Petr
1995

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Possibilistic Logic as Interpretability Logic

Petr Hájek

Technical report No. 630

February 1995

Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic
phone: (+422) 66414244 fax: (+422) 8585789
e-mail: hajek@uivt.cas.cz

INSTITUTE OF COMPUTER SCIENCE

ACADEMY OF SCIENCES OF THE CZECH REPUBLIC

Possibilistic Logic as Interpretability Logic

Petr Hájek [1] [2]

Technical report No. 630
February 1995

Abstract

It is shown that a variant of qualitative (comparative) possibilistic logic is closely related to modal interpretability logic, as studied in the metamathematics of first-order arithmetic. This contributes to our knowledge on the relations of logics of uncertainty to classical systems of modal logic.

Keywords

Uncertainty, Possibilistic logic, Interpretability logic.

[1] To appear in ADVANCES IN INTELLIGENT COMPUTING, edited by B.Bouchon-Meunier, R.R.Yager and L.A.Zadeh, Springer Verlag, 1995.

[2] Partial support by the COPERNICUS grant No. 10053 (MUM) is acknowledged.

1 Introduction

Possibilistic logic, as developed by Zadeh, Dubois, Prade and others (see e.g. [4] or [5]), deals with formulas and their possibilities, the possibility $\Pi(A)$ of a formula A being a real number from the unit interval, and the following axioms are assumed: $\Pi(\text{true}) = 1$, $\Pi(\text{false}) = 0$, equivalent formulas have equal possibilities, $\Pi(A \vee B) = max(\Pi(A), \Pi(B))$. It is very natural to ask how possibilistic logic relates to known systems of modal logics. This question was discussed in [1, 3, 6, 7]; in the last paper, possibilistic logic was related to tense (temporal) logic with finite linearly preordered time. One deals with Kripke models $\langle W, \Vdash \pi\rangle$ where W is a finite non-empty set of possible worlds, $\Vdash$ maps $Atoms \times W$ into $\{0, 1\}$ (truth evaluation), and π maps W into the unit interval [0,1]; sets $X \subseteq W$ have possibilities $\Pi(X) = max\{\pi(w) \mid w \in X\}$ and the possibility $\Pi(A)$ of a formula A is the possibility of the set of all worlds satisfying A. The mentioned three papers study the binary modality $\lhd$ defined as follows: $A \lhd B$ iff $\Pi(A) \leq \Pi(B)$. Classical Kripke models have the form $\langle W, \Vdash, R\rangle$ where R is a binary relation. In particular, each possibilistic model $\langle W, \Vdash, \pi\rangle$ determines a model $\langle W, \Vdash, R\rangle$ where $w_1 R w_2$ iff $\pi(w_1) \leq \pi(w_2)$; clearly, R is a linear preorder. [7] formulate an axiom system QPL sound for this semantics. The axioms are tautologies, transitivity $((A \lhd B)\&(B \lhd C)) \rightarrow (A \lhd C)$, linearity $(A \lhd B) \vee (B \lhd A)$, monotonicity $(A \lhd B) \rightarrow (A \vee C \lhd B \vee C)$; $0 \lhd A$ and $\neg(1 \lhd 0)$ (non-triviality); 1 is $true$. Deduction rules are modus ponens and the following necessitation: from $A \rightarrow B$ infer $A \lhd B$. It was proved in [1] that this axiom system is incomplete (even if complete for formulas with non-nested modalities) and an axiom scheme was exhibited making the system complete. A suggestion of Herzig has lead to the observation that the following pair of axioms sufficies:

(P) $\quad (A \lhd B) \rightarrow \Box(A \lhd B)$,

(P^-) $\quad \neg(A \lhd B) \rightarrow \Box\neg(A \lhd B)$,

where $\Box C$ is $\neg C \lhd 0$. The related system of tense logic of [1] has three necessity-like modalities G, H, I (meaning "in all future worlds", "in all past worlds", "in all present worlds", respectively); $\Box A$ is $HA\&IA\&GA$ and $A \lhd B$ is defined as $\Box(A \rightarrow \neg(I(\neg B)\&G(\neg B))$; equivalently, $\Box(A \rightarrow (JB \vee FB))$ where J is $\neg I \neg$ and F is $\neg G \neg$ (dual modalities). Details will not be repeated here.

In this paper, we are going to relate possibilistic logic to *interpretability logic*, as developed by Smoryński, Hájek, Švejdar, de Jongh, Veltman, Visser and others. Interpretability logic extends *provability logic* L, and we comment first on the latter. In provability logic, necessity (box, $\Box$) is understood as provability in a fixed axiomatic arithmetic T (e.g. Peano arithmetic). As Gödel discovered, in T we can define a formula $Pr(x)$ formalizing the notion of provability in T, e.g. $\neg Pr$ (false) is the formula Con expressing the consistency of T in T. (Gödel's second incompleteness theorem says that under reasonable assumptions on T, T does not prove its own consistency, i.e. $T \nvdash Con$.) Gödel also invented the method of self-reference in arithmetic, by constucting a formula ν such that $T \vdash \nu \equiv \neg Pr(\bar{\nu})$ (ν says "I am unprovable"; $\bar{\nu}$ is the numerical code of ν) and showed that under reasonable assumptions on T, ν is an independent formula ($T \nvdash \nu, T \nvdash \neg\nu$). (This is Gödel's first incompleteness theorem). An *arithmetical translation* of modal logic is a mapping $*$ associating with each formula

A of propositional modal logic (whose only modality is $\Box$) a sentence A^* of T in such a way that $*$ commutes with connectives (e.g. $(A\&B)^*$ is $A^*\&B^*$, etc.) and $(\Box A)^*$ is $Pr(\bar{A}^*)$ (this is how necessity is understood as provability). The *arithmetical completeness theorem* (cf. [12, 13]) says that a propositional modal formula A is provable in the provability logic L iff for each arithmetical translation $*$, $T \vdash A^*$. Provability logic also has its Kripke semantics and there is a corresponding completeness theorem (the same references).

For general theories T_1, T_2, T_1 is *interpretable* in T_2 if primitive notions of T_1 can be defined on T_2 in such a way that axioms of T_1 become provable in T_2. Consider the extension of our arithmetic T : $T_1 = (T + \phi)$, $T_2 = (T + \psi)$. In this case we can formalize the notion of interpretability, i.e., produce a formula $Intp(x, y)$ of T saying "$(T + \phi)$ interprets $(T + \psi)$" or "$(T + \psi)$ is interpretable in $(T + \phi)$"; T proves reasonable properties of this notion. This leads to a modal propositional logic with one unary modality $\Box$ and one binary modality $\lhd$; one has arithmetical interpretations ($(\Box A^*)$ is $Pr(\bar{A}^*)$, $(A \lhd B)^*$ is $Intp(\bar{A}^*, \bar{B}^*)$), arithmetical completeness, Kripke models, Kripke-style completeness [2, 8, 9, 10, 11, 14]. The double semantics of provability and interpretability logic (arithmetical and Kripke-like) gives interpretability logic its beauty; nevertheless, arithmetical interpretations will be disregarded here. We relate interpretability logic to a variant of qualitative possibilistic logic which we call the logic of sufficiently big possibilities (or the logic of future possibilities). In Section 1 we survey most basic facts on interpretability logic; in Section 2 we introduce our comparative logic of sufficiently big possibilities, in Section 3 we develop a tense logic with finite linearly preordered time and relate it to both preceding systems. Section 4 contains some remarks and Section 5 is an appendix giving indications for a completeness proof.

2 Preliminaries: Interpretability logics

Axioms are as follows:

axioms of L:
(L1) tautologies,
(L2) $\Box(A \to B) \to (\Box A \to \Box B)$,
(L3) $\Box A \to \Box\Box A$,
(L4) $\Box(\Box A \to A) \to \Box A$ (Löb's axiom),
additional axioms:
(J1) $\Box(A \to B) \to A \lhd B$
(J2) $((A \lhd B)\&(B \lhd C)) \to (A \lhd C)$
(J3) $((A \lhd C)\&(B \lhd C)) \lhd ((A \vee B) \lhd C)$
(J4) $A \lhd B \to (\Diamond A \to \Diamond B)$
(J5) $\Diamond A \lhd A$
(M) $(A \lhd B) \to ((A\&\Box C) \lhd (B\&\Box C))$
(P) $(A \lhd B) \to \Box(A \lhd B)$

Deduction rules: modus ponens and necessitation: from A infer $\Box A$. (Clearly, $\Diamond A$ is $\neg\Box\neg A$).

Remark Arithmetical validity of most axioms is easy to see; we comment on Löb's axiom. In fact, this is a variant of Gödel's second incompleteness theorem: by trivial manipulations, it can be written as $\Diamond\neg A \rightarrow \Diamond(\neg A \& \neg\Diamond\neg A)$, thus (replacing $\neg A$ by B) $\Diamond B \rightarrow \Diamond(B \& \neg\Diamond B)$ and hence $\Diamond B \rightarrow \neg\Box(B \rightarrow \Diamond B)$ which has the following arithmetical interpretation: if B is consistent (with T) then the formula $B \rightarrow Con(B)$ is unprovable (in T, thus: if $(T+B)$ is consistent then $(T+B)$ does not prove its own consistency; we disregard technical details.).

The following are important *axiom systems*: $IL = L + (J1) - (J5)$, $ILM = IL + (M)$, $ILP = IL + (P)$. See [2, 8, 9, 11, 14].

Note that in IL, box is definable from triangle:

$$IL \vdash \Box A \equiv (\neg A \lhd 0)$$

where 0 is *false*). A *Veltman model* has the form $\langle W, \Vdash, R, S\rangle$ where $\langle W, \Vdash, R\rangle$ is a Kripke model with R transitive and asymmetric (hence irreflexive) and S is a reflexive transitive relation containing R. (This is only a particular case; see [10] for the general case). One defines $w \Vdash \Box A$ iff for all $v \in W$, wRv implies $v \Vdash A$; $w \Vdash A \lhd B$ iff

$$(\forall v)(wRv \,\&\, v \Vdash A \;\rightarrow\; (\exists u)(wRu \,\&\, vSu \,\&\, u \Vdash B)).$$

The *completeness theorem* for ILP says that $ILP \vdash A$ iff A is true in each finite Veltman model $\langle W, \Vdash, R, S\rangle$ satisfying the following condition:

$$(wRv \,\&\, wRu \,\&\, vSu \,\&\, wRw' \,\&\, w'Rv) \rightarrow w'Ru$$

Another formulation is as follows: let vS_wu mean $wRv \,\&\, wRu \,\&\, vSu$. Then

$$(vS_wu \,\&\, wRw' \,\&\, w'Rv) \rightarrow vS_{w'}u.$$

To get completeness for IL and ILM one needs a more complicated notion of a Veltman model.

3 The comparative logic of future possibilities

Comparing ILP with $QPL + (P)$ we see that $QPL + (P)$ proves (J1-J4) but not (J5) and clearly does not prove Löb's axiom. If one restricts oneself to positive models (for each $w \in W, \pi(w) > 0$) then $\Box$ of QPL becomes an (S5)-modality; in particular, $\Box A \rightarrow A$ is sound; $QPL + (P) + (\Box A \rightarrow A)$ axiomatizes completely $\lhd$ with respect to positive models.

Our aim is to relate possibilistic logic more closely to interpretability logic. This is done below.

To marry ILP with possibility theory, consider the world-dependent *future possibility*: $\Pi(A, w) = sup\{\pi(w') \mid w' > w \text{ and } w' \Vdash A\}$. Here $w' > w$ means $\pi(w') > \pi(w)$. Define $w \Vdash A \lhd B$ if $\Pi(A, w) \leq \Pi(B, w)$. Thus $A \lhd B$ is satisfied in the world w if either $\Pi(A)$, $\Pi(B) \leq \pi(w)$ or $\Pi(A) \leq \Pi(B)$. This suggests the following (fuzzy)

reading of the new triangle-modality: the possibility of A is less-than-or-equal to the possibility of B, or neither A nor B are too much possible.

Our comparative *logic of future possibilities* (or, if the reader prefers, *logic of sufficiently big possibilities*) has formulas built from propositional variables using connectives and the modality $\lhd$; its models are finite possiblistic Kripke models $K = \langle W, \Vdash -, \pi, \rangle$ and the semantics of $\lhd$ is given by comparison of future possibilities as above. We shall find a complete axiomatization.

Define $\Box$ to be the *future necessity*: $w \Vdash -\Box A$ iff for all $w' > w$, $w' \Vdash -A$. This relates possibilistic logic and its Kripke models to tense logic and its Kripke models; we shall investigate the corresponding tense logic in the next section. At this moment, let us stress that relating possibilistic logic to tense logics (and other logics, e.g. interpretability logic) should contribute to our understanding of what possibility theory is; one interpretation is that $\Pi(A)$ is the *last moment* in which A is possible (or zero). This temporal interpretation makes our future possibility natural: $\Pi(A, w)$ means (in the world w) the *last moment after now* in which A is possible (or zero). To elucidate this, let us verify the validity of the axiom (J5): $\Diamond A \lhd A$. Given w, if $\Pi(\Diamond A, w) > 0$, let w' be the last world after w satisfying $\Diamond A$; thus there is a last $w'' > w$ satisfying A; $\Pi(\Diamond A, w) = \Pi(w') < \Pi(w'') = \Pi(A, w)$. Clearly, each possibilistic model $\langle W, \Vdash -, \pi \rangle$ determines a particular Veltman model called an LPO-model (linear preorder):

An *LPO-model* is a Veltman model $\langle W, \Vdash -, R, S \rangle$ where S is a linear preorder of W, i.e. S is transitive and dichotomous (wSv or vSw for all $w, v \in W$) and R is the corresponding strict preorder: wRv if wSv and not vSw.

Observe that $w \Vdash -A \lhd B$ in the possibilistic model $\langle W, \Vdash -, \pi \rangle$ (with respect to the future-possibility semantics) iff $w \Vdash -A \lhd B$ in the corresponding LPO model $\langle W, \Vdash -, R, S \rangle$ (with respect to Veltman semantics). Let (D) be the axiom of dichotomy $(A \lhd B) \vee (B \lhd A)$. Then:

Fact Axioms of ILPD are tautologies of LPO-models.

But we shall show that ILPD is not complete for LPO-models. In the sequel we develop a tense logic with finite linearly preordered time and one (future) necessity extending ILPD and complete for LPO -models.

4 A tense logic with finite linearly preordered time

In this section we shall investigate the modality of future necessity introduced above. Note that here we have only the future necessity (always in the future), no past necessity and no present necessity. This is in contrast to the system of [1] discussed in Section 1, with the same Kripke models (with finite linearly preordered time) but with three necessities mentioned. We show that our present tense logic is completely axiomatized by an axiom system extending the axioms of provability logic L by a single axiom of linear preorder:

Let (E) be the axiom

$$\Box(\Box A \to B) \vee \Box(\Box B \to \Box A)$$

(linEar prEordEr).

Theorem 4.1 *1. The logic $L + (E)$ (provability logic plus (E)) is sound and complete for finite Kripke models $\langle W, \Vdash, R\rangle$ such that there is a linear preorder S on W (transitive and dichotomous) whose corresponding strict preorder is R (thus wRv iff wSv and not vSw; in other words, $\langle W, \Vdash, R, S\rangle$ is an LPO-model).*

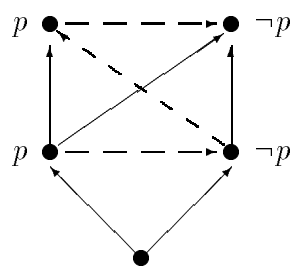
2. In $L + (E)$ define $\lhd$ as follows (definition):

(F) $A \lhd B$ iff $\Box(\neg A \& \neg B) \vee \Diamond(B \& \Box\neg A)$.

Then $L + (E) + (F)$ proves all axioms of ILPD.

3. The definition (F) is true in each LPO-model if $\lhd$ means comparison of future possibilities.

4. ILPD does not prove (E); in particular, the model



is a model of ILPD, but the formula $\Box(\Box p \to \neg p) \vee \Box(\Box\neg p \to \Box p)$ is false in the root.

The *proof* of (1) is a modification of the proof of the completeness theorem in [1]; (2)-(4) are easy.

Summarizing, $L+(E,F)$ is the logic of comparison of future possibilities (over finite models); it strictly extends the interpretability logic $ILPD$.

5 Remarks

We continue with a series of remarks.

Remark Over LPO models, $\Box$ and $\lhd$ are interdefinable: (F) defines $\lhd$ from $\Box$ and evidently $\Box A \equiv (\neg A \lhd 0)$ is a tautology. Thus the theory L + (E,F) may be presented as a theory with a single modality $\lhd$.

Remark We compare the modality of comparison of future possibilities (denoted $\lhd_f$ in the present remark) with the (world independent) modality of comparison of possibilities ($\lhd_c, c$ for "constant over worlds"). For a moment, let a CNN-formula (closed non nested) be a boolean combination of formulas of the form $A \lhd B$. If C is such a formula, C_f and C_c mean results of replacement of $\lhd$ by $\lhd_f$ and $\lhd_c$ respectively. The relation is as follows:

Fact Let C be a CNN-formula. If C_f is a tautology (over LPO-models) then C_c is a tautology. On the other hand, for each axiom C of QPL except $\neg(1 \lhd 0), C_f$ is a tautology. (But $1 \lhd_f 0$ is satisfied in each maximal element of each LPO-model).

The first part is proved by adding to an arbitrary LPO-model a new least element.

Let QPL_0 result from QPL by removing the axiom $\neg(1 \lhd 0)$. It can be shown, in an analogy to [1], that QPL_0 axiomatizes all CNN formulas that are tautologies in the future semantics. Alternatively, it axiomatizes all CNN formulas that are tautologies in the constant semantics with respect to all LPO-models plus the *empty* model (no worlds, possibility of each formula is 0).

Remark The task remains to analyze the meaning of ILPD (which is weaker than $L + (E, F)$) for possibilistic logic. As the example above shows, ILPD admits models substantially different from (not elementarily equivalent to) any LPO-model. Does every model of ILPD have a possibilistic interpretation?

Remark Studying future possibility we restricted ourselves to finite models. Such a restriction is fully justified in the case of constant possibilities, as it was shown in [1]. Here this is an additional assumption, justified, e.g. by postulating that the scale of possibilities ($\pi(w)$) is a finite subset of [0,1]. Expressive power of future possibility comparison on infinite models might deserve additional consideration.

6 Appendix

A proof of the completeness theorem for $L + (E)$ may be obtained by an inspection of the proof of completeness of FLPOT in [1] Section 3. We briefly indicate the necessary changes; the reader is assumed to have a copy of [1] at his/her disposal.

In 3.5, change the definition of E:

$\Gamma_1 E \Gamma_2$ iff, for each C, $GC \in \Gamma_1$ iff $GC \in \Gamma_2$

(G replaces $\Box$; we have no H, I modalities). Delete 3.6, 3.7, 3.8(1) holds as it stands; 3.8(2) will read: If $\Gamma E \Gamma'$ then $\Gamma R \Delta$ implies $\Gamma' R \Delta$ (and also $\Delta R \Gamma$ implies $\Delta R \Gamma'$, but this will be proved later).

Proof: If $GA \in \Gamma'$ then $GA \in \Gamma$ and $A \in \Delta$.

Define irreflexive theories as in 3.9; 3.10 holds as it stands.
3.11(1) holds by a new proof:

Proof: Assume $\Theta R \Gamma_1$, $\Theta R \Gamma_2$, and let all of the following three conditions be false: $\Gamma_1 R \Gamma_2$, $\Gamma_2 R \Gamma_1$, $\Gamma_1 E \Gamma_2$. Then there are A, B, C witnessing this, i.e.

$$(GA \& \neg B \& GC) \in \Gamma_1$$

$$(GB \& A \& \neg GC) \in \Gamma_2$$

But $\Theta \vdash G(G(A\&C) \to B) \vee G(GB \to G(A\&C))$, thus Θ proves one of the disjuncts. If $\Theta \vdash G(G(A\&C) \to B)$ then $(G(A\&C) \to B) \in \Gamma_1$ and Γ_1 is inconsistent; if $\Theta \vdash G(GB \to G(A\&C)$ then $\Gamma_2 \vdash GB \to G(A\&C) \to GC$, thus Γ_2 is inconsistent. But Γ_1, Γ_2 are consistent.

In Def.3.12 of a critical formula delete H, I; define $S = \{\Gamma \mid \Delta R\Gamma \text{ or } \Delta E\Gamma\}$. 3.13 is O.K.

In 3.14 show that, if Ω is a set of irreflexive theories, then for $\Gamma_1, \Gamma_2 \in \Omega$ *exactly* one of the conditions $\Gamma_1 R \Gamma_2$, $\Gamma_1 E \Gamma_2$, $\Gamma_2 R \Gamma_1$ holds; i.e., that they are mutually exclusive; use the proof is it stands. For transitivity, we have to prove the second half of 3.8:

$$\Gamma_1 R \ \ \Gamma_2 E \ \ \Gamma_3 \text{ implies } \Gamma_1 R \ \ \Gamma_3.$$

If $\Gamma_1 R \ \ \Gamma_2 E \ \ \Gamma_3$ and not $\Gamma_1 R \ \ \Gamma_3$ then $\Gamma_3 R \ \ \Gamma_1$ or $\Gamma_1 E \ \ \Gamma_3$. But $\Gamma_3 R \ \ \Gamma_1$ implies $\Gamma_3 R \ \ \Gamma_2$ which is incompatible with $\Gamma_2 E \ \ \Gamma_3$, and $\Gamma_1 E \ \ \Gamma_3$ gives $\Gamma_1 E \ \ \Gamma_2$ which is incompatible with $\Gamma_1 R \ \ \Gamma_2$.

3.15(1) is O.K. as its stands and (2), (3) are deleted. Similarly for 3.16(1), (2), (3). This completes the proof; the rest is O.K.

Bibliography

[1] Bendová K., Hájek P.: Possibilistic logic as tense logic, in: Proc. QUARDET'93, Barcelona 1993.

[2] Berarducci A.: The interpretability logic of Peano arithmetic, Journal of Symb. Log. 55, 1990, 1059-1089.

[3] Boutilier C.: Modal logics for qualitative possibility and beliefs, in: Uncertainty in Artificial Intelligence VIII, Morgan-Kaufmann Publ. 1992, 17-24.

[4] Dubois D., Lang J., Prade H.: Possibilistic logic, Report IRIT-91-98/R, Decembre 1991, Toulouse.

[5] Dubois D., Prade H.: Possibilistic and fuzzy logic, in: (Smets et al., eds): Non-standard logics for automated reasoning, Academic Press 1988, 287.

[6] Dubois D., Prade H., Testemate C.: In search for a modal system for possibility theory, in: Proc. ECAI'89, Munich 1988, 581-586.

[7] Fariñas del Cerro L., Herzig A.: A modal analysis of possibilistic logic, in: (Kruse et al.,eds): Symbolic and Quantitative Approaches to Uncertainty, Lect. Notes in Comp. Sci. 548, Springer-Verlag 1991, p.58, (cf. also: Jorrand and Kelemen, eds: Fundamentals of AI research, Lect. Notes in AI 535, Springer-Verlag 1991, 11).

[8] Hájek P., Montagna F.: The logic of Π_1 - conservativity, Arch. Math. Logic 30, 1990, 113-123.

[9] Hájek P., Montagna F.: The logic of Π_1 - conservativity continued, Arch. Math. Logic 32, 1992, 57-63.

[10] de Jongh D., Veltman F.: Provability logics for relative interpretability, in: Mathematical Logic (Proceedings of the Heyting conference, Bulgaria 1988), Plenum Press 1990.

[11] Shavrukov V.Yu.: The logic of relation interpretability over Peano arithmetic (in Russian) Moscow 1988, preprint.

[12] Smoryński C.: Self-reference and modal logic, Springer-Verlag 1988.

[13] Solovay R.: Provability interpretations of modal logic, Israel Journ. Math. 25, 1976, 287-304.

[14] Visser A.: Interpretability logic, in: Mathematical Logic (Proc. Heyting conference Bulgaria 1988), Plenum Press 1988.