Heinrich Wansing

A General Possible Worlds Framework for Reasoning about Knowledge and Belief

Abstract. In this paper non-normal worlds semantics is presented as a basic, general, and unifying approach to epistemic logic. The semantical framework of non-normal worlds is compared to the model theories of several logics for knowledge and belief that were recently developed in Artificial Intelligence (AI). It is shown that every model for implicit and explicit belief (Levesque), for awareness, general awareness, and local reasoning (Fagin and Halpern), and for awareness and principles (van der Hoek and Meyer) induces a non-normal worlds model validating precisely the same formulas (of the language in question).

1. Introduction

Given the standard Kripkean model theory of normal modal propositional logic, the basic idea of Hintikka-style epistemic logic (cf. [Hil]) is to make use of the analogy between modal notions ("it is neccessary that") and epistemic notions ("it is known that"). Under the epistemic interpretation of the necessity operator $\Box$, the normal modal propositional logic $KT45$ ($= S5$) becomes the standard logic of *knowledge*. Usually the 'knowledge' of a knowledge base is, however, not required to be true. In this case the weaker logic $KD45$ is known as the standard logic of *belief.* The general methodological conception behind this approach is that particular interpretations of modal operators require particular modal axioms (or, in semantical terms, particular algebraic properties of the binary 'accessibility' relation R among possible worlds). Taking agents $1, \ldots, n$, i.e. subjects of knowledge and belief, into account, one arrives at multi-modal logic, disposing of modal operators $\Box_1, \ldots, \Box_n$ and binary relations $R_1, \ldots, R_n$.

Yet, from the very beginning of epistemic logic it was quite obvious that the model theory of normal modal propositional logic could not fully do justice to the 'hyper-intensional' ([Cr2]) character of the 'subjective modalities' of knowledge and belief. It was generally accepted that $KT45$ and $KD45$ merely captured highly idealised notions of knowing and believing. Put to the test of thorough philosophical examination (see e.g. [Sa]), it seems very doubtful indeed, if there *at all exists* something like a *logic* of knowledge and belief, being a general logic of these two *notions.* An agent's knowledge or belief, it seems, might in fact be completely irregular. In [Ra1], [Ra2], it is suggested that "knows that" and "believes that" should therefore best be regarded as *descriptive* rather than *logical* expressions. The idea that knowledge and belief are descriptive notions will in the sequel act as a kind of guiding thought. The resulting strategy is to look for a *basic and general semantical framework* that (i) is flexible enough to interpret attitude verbs like "knows that" and "believes that" as descriptive expressions, and (ii) admits of an extension to modelling the 'limiting cases' of knowledge and belief, (multi-modal) $KT45$ and $KD45$. It will be argued that non-normal worlds semantics as developed in [Ra1], [Ra2], [Wa], [PW] does constitute such a general framework.[1]

In recent years, research in Artificial Intelligence (AI) has breathed new life into epistemic logic. Due to a certain stagnation of ordinary Hintikka-style epistemic logic, standard methodological principles and the pre-theoretical notions of knowledge and belief were re-thought and sophisticated modifications as well as extended formalisms developed, cf. e.g. [Le], [Va], [FH], [vdHM]. Although these authors advocate a pragmatic attitude towards epistemic logic, they have so far not presented any basic, unifying framework. Moreover, AI researchers more or less unanimously reject the non-normal worlds approach to knowledge and belief. In this paper it will be shown, however, that the knowledge and belief structures in [Le], [Va], [FH], [vdHM] can be viewed as special versions of non-normal worlds semantics.

2. Possible worlds and 'logical omniscience'

It is well known that multi-modal Kripke models $\mathcal{M} = \langle W, R_1, \ldots, R_n, V\rangle$ give rise to 'logical omniscience' ([Hi2]), i.e. closure properties such as:

- if φ is a classical tautology, then $\mathcal{M} \models \Box_i \varphi$ (belief of classical tautologies),
- if $\mathcal{M}, w \models \Box_i \varphi$, $\mathcal{M}, w \models \Box_i(\varphi \to \psi)$, then $\mathcal{M}, w \models \Box_i \psi$ (closure under implication),
- if $\mathcal{M}, w \models \Box_i \varphi$, $\mathcal{M} \models \varphi \to \psi$, then $\mathcal{M}, w \models \Box_i \psi$ (closure under valid implication).

Some authors (cf. e.g. [Pa], [Ha1]) come close to claiming that the possible worlds paradigm itself is committed to the above closure properties, but this is not quite correct. Belief of classical tautologies is, of course, avoided by e.g. taking an intuitionistic modal base, and closure under valid implication is e.g. no by-product of the neighbourhood semantics for classical modal logics (cf. e.g. [Ch]). However, believing or knowing all theorems of Heyting's propositional logic or some other logical system is still a form of 'logical omniscience' just as in classical modal logic the closure condition:

- if $\mathcal{M}, w \models \Box_i \varphi$, $\mathcal{M} \models \varphi \leftrightarrow \psi$, then $\mathcal{M}, w \models \Box_i \psi$ (substitutivity of equivalents).[2]

[1] Very much the same ideas have also been developed by Cresswell, cf. e.g. [Cr1].

[2] In general, turning to subsystems may even increase computational complexity.

Substitutivity of equivalents plays a prominent role in the philosophy of possible worlds. If *propositions* are introduced as sets of possible worlds, the problem with substitutivity of equivalents naturally leads to a search for *more fine-grained propositions.* At this point notions of intensional isomorphism ([Ca]) etc. come into play. Whatever intrinsic interest notions like these deserve, I think that in the case of knowledge and belief they somewhat miss the point. Belief and knowledge are 'subjective' notions with strong psychological components.[3] It is conceivable that an agent's knowledge and belief does not display even the weakest regularities describable in terms of inference patterns. Moreover, intuitively an agent's limited reasoning capacities cannot necessarily be expected to conform to some sort of closure property. Still, even if there is no *general logic* of knowledge and belief, epistemic logic is, of course, not merely concerned with the descriptive task of modelling some agent's factual knowledge or belief. For purposes of design there arises a variety of (restricted) regularity requirements. From a methodological point of view, it would be desirable to meet such requirements by imposing certain restrictions on a basic descriptive framework. Now, as pointed out in [Hi2], logical omniscience does not arise, if non-normal ("impossible possible") worlds are permitted. In what follows, non-normal worlds semantics will be offered as a general possible worlds framework in which knowledge and belief may be treated as something like *primitive intensional operators.* In this framework, degrees of freedom are provided by various constraints on admissible valuations and binary 'accessibility' relations among worlds.

3. The framework of non-normal worlds semantics

If "knows that" and "believes that" are viewed as descriptive expressions, one might think of an analogy to the interpretation of descriptive expressions in first-order logic. The semantics of first-order logic does not account for any particular deductive behaviour of a non-logical relation symbol Q. Instead, principles governing Q are explicitly stated in the form of a first-order theory $Th(Q)$, and attention may then be restricted to $Mod(Th(Q))$, i.e. the class of all first-order models of $Th(Q)$. Now, we are interested in a semantics for knowledge and belief that is not bound to validate any particular axioms or inference rules involving $\Box_i$. The class of all knowledge or belief structures should characterise just the background-logic in question, here classical propositional logic. Regularity requirements are then to be explicitly stated, either in the form of modal axioms like $(\Box_i \varphi \wedge \Box_i \psi) \rightarrow \Box_i(\varphi \wedge \psi)$ or in the form of modal inference rules like

---

[3] There is, of course, also an 'objective' notion of knowledge, that may be read as "having in principle the relevant information available". But, as [Ba, p. 4] tells us, "[i] information travels at the speed of logic, genuine knowledge only travels at the speed of cognition and inference".

- $\vdash \varphi \leftrightarrow \psi \Rightarrow \vdash \Box_i \varphi \leftrightarrow \Box_i \psi$
  (congruence),
- $\vdash \varphi \Rightarrow \vdash \Box_i \varphi$
  (necessitation).

Attention may then be restricted to the appropriate classes of knowledge or belief structures.

The main idea of Rantala's non-normal worlds semantics ([Ra1], [Ra2]) is to allow for reference points which do not impose any compositional structure on the evaluation of formulas. If some such non-standard world w^* is involved in the evaluation of a modal formula $\Box_i \varphi$ at some world w, the syntactic structure of φ and the truth of particular other formulas at w^* need not affect the truth or falsity of $\Box_i \varphi$ at w. Let $\mathbf{L}$ be the language of multi-modal propositional logic. Starting with ordinary Kripke models, one obtains *Rantala models* $\mathcal{M} = \langle W, W^*, R_1, \ldots, R_n, V \rangle$ where W, W^* are sets (of 'normal' and 'non-normal' worlds respectively), $W \neq \emptyset$, $R_i (1 \leqslant i \leqslant n)$ is a binary relation on $W \cup W^*$, and V: $FORM_{\mathbf{L}} \times (W \cup W^*) \to \{0, 1\}$ such that:

- $\forall w \in W$,
  $V(\varphi \wedge \psi, w) = 1 \Leftrightarrow V(\varphi, w) = V(\psi, w) = 1$,
  $V(\neg \varphi, w) = 1 \Leftrightarrow V(\varphi, w) = 0$,
  $V(\Box_i \varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)(R_i ww' \Rightarrow V(\varphi, w') = 1)$.

In particular, the truth or falsity of formulas need not be *recursively specified* at non-normal worlds. An $\mathbf{L}$-formula φ is true in $\mathcal{M} = \langle W, W^*, R_1, \ldots, R_n, V \rangle$ at $w \in W \cup W^*$ ($\mathcal{M}, w \models \varphi$) iff $V(\varphi, w) = 1$. Validity is defined in terms of *normal worlds*: φ is *valid in* $\mathcal{M} = \langle W, W^*, R_1, \ldots, R_n, V \rangle$ ($\mathcal{M} \models \varphi$) iff $\forall w \in W$, $V(\varphi, w) = 1$; φ is *valid* simpliciter ($\models \varphi$) iff for every Rantala model $\mathcal{M}$, $\mathcal{M} \models \varphi$. Of course, non-normal worlds enter the picture by way of the key clause for evaluating modal formulas at normal worlds. Distinguishing among different types of valuations, one obtains a *generalised notion of completeness* wrt Rantala frames, i.e. structures $\langle W, W^*, R_1, \ldots, R_n \rangle$ (cf. [Ra2], [Wa], [PW]). Let $\mathcal{S}$ be a modal system, VAL be the class of all valuations (according to the definition of Rantala models), and $U \subseteq VAL$. Call a Rantala model $\langle W, W^*, R_1, \ldots, R_n, V \rangle$ a *U-model* iff $V \in U$. $\mathcal{S}$ is said to be *U-determined by a class of Rantala frames* $\mathbf{C}$ iff for every $\varphi \in \mathbf{L}$ the following holds: $\vdash_{\mathcal{S}} \varphi$ iff for every frame $\mathcal{F} \in \mathbf{C}$ and every U-model $\mathcal{M}$ based on $\mathcal{F}$, $\mathcal{M} \models \varphi$. $\mathcal{S}$ is *incomplete* iff there is no valuation-class $U \subseteq VAL$ and no class of Rantala frames $\mathbf{C}$ such that $\mathcal{S}$ is U-determined by $\mathbf{C}$. This generalised notion of completeness has 'empirical content': there are normal modal propositional logics which are incomplete wrt the Rantala semantics ([Wa], [PW]).

Does this semantics in fact constitute a basic, general, and possibly unifying framework for epistemic logic? As pointed out (for the one-agent case) in [Wa], [PW], the requirement of being *basic* is met:

FACT 1. *Classical propositional logic is VAL-determined by the class of all Rantala frames.*

Concerning *generality*, one may notice a division of labour between the available degrees of freedom. Regularity requirements in terms of inference rules are captured by suitable choices of valuation types; modal axioms are typically taken into account by properties of the relations R_i. 'Typically', because the validity in a Rantala model of the K-axiom $\Box_i(\varphi \to \psi) \to \to (\Box_i \varphi \to \Box_i \psi)$ calls for no relational constraint but for a semantic version of modus ponens in non-normal worlds:

(†) $(\forall w^* \in W^*)$ if $(V(\varphi, w^*) = V(\varphi \to \psi, w^*) = 1)$, then $V(\psi, w^*) = 1$.

The valuation types CGR and NEC are defined as follows:

- CGR is the biggest $X \subseteq VAL$ such that $\forall \varphi, \psi \in \mathbf{L}$, $\forall V \in X$, if $((\forall w \in W)\, V(\varphi, w) = V(\psi, w))$, then $(\forall w^* \in W^*)$ if $((\exists w \in W)\, R_i ww^*)$, then $V(\varphi, w^*) = V(\psi, w^*)$;
- NEC is the biggest $X \subseteq VAL$ such that $\forall \varphi, \psi \in \mathbf{L}$, $\forall V \in X$, (†) and if $(\forall w \in W)\, V(\varphi, w) = 1$, then $(\forall w^* \in W \cup W^*)$ if $((\exists w \in W)\; R_i ww^*)$, then $V(\varphi, w^*) = 1$.

FACT 2. i) *The (multi-modal) minimal classical modal logic E is CGR-determined by the class of all Rantala frames.*

ii) *The (multi-modal) minimal normal modal logic K is NEC-determined by the class of all Rantala frames.*[4]

It is almost trivially true that the 'limiting cases' (multi-modal) $KD45$ and $KT45$ can be handled within non-normal worlds semantics, because ordinary Kripke models *are* Rantala models. But also:

FACT 3. *(Multi-modal) $KD45$ ($KT45$) is NEC-determined by the class of all Rantala frames $\langle W, W^*, R_i, \ldots, R_n \rangle$, in which $R_i (1 \leqslant i \leqslant n)$ is serial, transitive, and Euclidean (is an equivalence relation) on $W \cup W^*$.*

In designing reasonable logics of knowledge and belief, comparatively weak closure properties might still turn out to be inappropriate because of being formulated in terms of arbitrary formulas. It might e.g. be useful to restrict particular modal inference rules to certain subsets of $FORM_{\mathbf{L}}$. Restricted modal rules of inference can be dealt with in non-normal worlds semantics in a completely straightforward and regular fashion. If an inference rule is restricted to a set $Y \subseteq FORM_{\mathbf{L}}$, one simply has to restrict the defining clause of the corresponding valuation type to Y (cf. [Wa]).

The present version of non-normal worlds semantics is still quite 'normal', since non-normal worlds are added to ordinary Kripke models for *normal* modal propositional logic. However, for this reason (normal) modalities and epistemic interpretations of them are still available, restricting their evaluation

[4] In [Wa], [PW] it is emphasised that these characterisations of 'minimal' modal logics display a certain uniformity of semantic modelling across different lattices of modal systems: in each case the class of *all* Rantala frames is involved.

clauses in (*NEC*-) models to normal worlds. Thus, descriptively adequate belief and logical necessity ('omniscient belief') can simultaneously be interpreted in Rantala models (cf. the model constructions of the following section). In what follows, inductive evidence is provided for the claim that non-normal worlds semantics also constitutes a *unifying* framework for reasoning about knowledge and belief.

4. Relating the model theories of some recent logics of belief to non-normal worlds semantics

In this section non-normal worlds semantics will be compared to the model theories of some of the most sophisticated logics of belief that are currently available: Levesque's logic of implicit and explicit belief ([Le]), Fagin and Halpern's logic of awareness, logic of general awareness, and logic of local reasoning ([FH]), and van der Hoek and Meyer's logic of awareness and principles ([vdHM]). All these logics partially overcome logical omniscience. In order to keep this paper self-contained and for the reader's convenience in each case a short description is given of the semantics in question. For details and motivations, however, the reader is referred to the original papers.

4.1. Levesque's logic of explicit and implicit belief

Levesque suggests looking at omniscient belief as what he calls *implicit belief*, where an agent's implicit belief follows from what is "actively held to be true" ([Le, p. 198]) by the agent, in other words, from what the agent *explicitly* believes. This distinction is reflected in the formal propositional language, say $\mathbf{L}_1$, considered by Levesque. In addition to the standard vocabulary of non-modal propositional logic, there are (unary) intensional operators L, B for (one-agent) implicit and explicit belief respectively. Nestings of intensional operators are not permitted, i.e., only *propositional formulas* may occur within the scope of L and B.

In his semantics, Levesque employs *partial* and also *incoherent* reference points. We have *structures for implicit and explicit belief* $\mathcal{M} = \langle S, \mathcal{B}, T, F\rangle$, where S is a (non-empty) set of *situations*, $\mathcal{B} \subseteq S$, and T, F are mappings from $PRIM$ (the set of all sentence letters) into 2^S. $T(p)$ is intuitively to be interpreted as the set of situations *supporting the truth* of p, $F(p)$ as the set of situations *supporting the falsity* of p. In particular *partial* situations s, where $s \notin T(p) \cup \cup F(p)$, and incoherent situations s, where $s \in T(p) \cap F(p)$ are permitted for some $p \in PRIM$; s is called a *possible world* iff for all $p \in PRIM$ ($s \in T(p)$ or $s \in F(p)$) but $s \notin T(p) \cap F(p)$. A possible world s is *compatible* with a situation s' if the following holds for all $p \in PRIM$: if $s' \in T(p)$, then $s \in T(p)$, and if $s' \in F(p)$, then $s \in F(p)$. Thus, possible worlds can be compatible only with coherent situations. Let $\mathcal{W}(\mathcal{B})$ be the set of all possible worlds in S compatible with some situation in $\mathcal{B}$. Working with a partial semantics, there is a distinction between two *support relations* $\models_T$ and $\models_F \subseteq S \times FORM_{\mathbf{L}_1}$. The notions $\mathcal{M}, s \models_T \varphi$ (situation s supports the truth of φ in $\mathcal{M}$), $\mathcal{M}, s \models_F \varphi$ (situation s supports the falsity of φ in $\mathcal{M}$) are inductively defined as follows:

- $\mathcal{M}, s \models_T p \Leftrightarrow s \in T(p)$,
  $\mathcal{M}, s \models_F p \Leftrightarrow s \in F(p)$;
- $\mathcal{M}, s \models_T \varphi \wedge \psi \Leftrightarrow \mathcal{M}, s \models_T \varphi$ and $\mathcal{M}, s \models_T \psi$,
  $\mathcal{M}, s \models_F \varphi \wedge \psi \Leftrightarrow \mathcal{M}, s \models_F \varphi$ or $\mathcal{M}, s \models_F \psi$;
- $\mathcal{M}, s \models_T \neg \varphi \Leftrightarrow \mathcal{M}, s \models_F \varphi$,
  $\mathcal{M}, s \models_F \neg \varphi \Leftrightarrow \mathcal{M}, s \models_T \varphi$;
- $\mathcal{M}, s \models_T B\varphi \Leftrightarrow (\forall t \in \mathcal{B}) \mathcal{M}, t \models_T \varphi$,
  $\mathcal{M}, s \models_F B\varphi \Leftrightarrow \mathcal{M}, s \not\models_T B\varphi$;
- $\mathcal{M}, s \models_T L\varphi \Leftrightarrow \big(\forall t \in \mathcal{W}(\mathcal{B})\big) \mathcal{M}, t \models_T \varphi$,
  $\mathcal{M}, s \models_F L\varphi \Leftrightarrow \mathcal{M}, s \not\models_T L\varphi$.

If s is a possible world, then φ is said to be *true at s* (*false at s*) iff $\mathcal{M}, s \models_T \varphi$ ($\mathcal{M}, s \models_F \varphi$). Also validity is defined in terms of possible worlds only; φ is *valid in* $\mathcal{M} = \langle S, \mathcal{B}, T, F \rangle$ ($\mathcal{M} \models \varphi$) iff φ is true at every possible world $s \in S$; φ is *valid* simpliciter ($\models \varphi$) iff for any structure for implicit and explicit belief $\langle S, \mathcal{B}, T, F \rangle$, φ is true at every possible world $s \in S$. Thus, if for $\mathcal{M} = \langle S, \mathcal{B}, T, F \rangle$ S contains no possible world at all, then $\mathcal{M}$ is a trivial model validating every $\mathbf{L}_1$-formula.

CLAIM 1. *Every nontrivial structure for implicit and explicit belief induces a Rantala model validating precisely the same* $\mathbf{L}_1$*-formulas.*

PROOF. Given a nontrivial structure for implicit and explicit belief $\mathcal{M} = \langle S, \mathcal{B}, T, F \rangle$, define a Rantala model $\mathcal{M}_R = \langle W, W^*, R_B, R_L, V \rangle$ as follows:

- $W =_{def} \{w \in S | (\forall p \in PRIM)\big(w \in T(p) \vee w \in F(p)\big) \wedge w \notin T(p) \cap F(p)\}$;
- $W^* =_{def} S \setminus W$;
- $R_B =_{def} \{\langle w, w' \rangle | w \in S, w' \in \mathcal{B}\}$;
- $R_L =_{def} \{\langle w, w' \rangle | w \in S, w' \in \mathcal{W}(\mathcal{B})\}$;
- V: $FORM_{\mathbf{L}_1} \times (W \cup W^*) \to \{0, 1\}$ such that:
  $\forall w \in W \cup W^*$,
  $V(p, w) = 1 \Leftrightarrow w \in T(p) \wedge w \notin T(p) \cap F(p)$,
  $V(p, w) = 0 \Leftrightarrow w \in F(p) \vee w \notin T(p) \cup F(p) \vee w \in T(p) \cap F(p)$,
  $\forall w \in W$,
  $V(\varphi \wedge \psi, w) = 1 \Leftrightarrow V(\varphi, w) = V(\psi, w) = 1$,
  $V(\neg \varphi, w) = 1 \Leftrightarrow V(\varphi, w) = 0$,
  $V(B\varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)\big(R_B ww' \Rightarrow V(\varphi, w') = 1\big)$,
  $V(L\varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)\big(R_L ww' \Rightarrow V(\varphi, w') = 1\big)$.

Note that by induction on $\varphi \in \mathbf{L}_1$ it can be shown that for every $w \in W$, i.e. every possible world in S, $\mathcal{M}, w \not\models_T \varphi \Leftrightarrow \mathcal{M} \models_F \varphi$. We will show by induction on $\varphi \in \mathbf{L}_1$ that $(\forall \varphi \in \mathbf{L}_1)$ $(\forall w \in W)$ $\mathscr{M}, w \models \varphi \Leftrightarrow \mathscr{M}_R, w \models \varphi$, whence the claim follows. Suppose that $w \in W$:

- $\mathscr{M}, w \models_T p \Leftrightarrow w \in T(p) \overset{def.}{\Leftrightarrow} w \in T(p) \wedge w \notin T(p) \cap F(p) \overset{def.}{\Leftrightarrow} \mathscr{M}_R, w \models p$;
- $\mathscr{M}, w \models_T \varphi \wedge \psi \Leftrightarrow \mathscr{M}, w \models_T \varphi$ and $\mathscr{M}, w \models_T \psi$ $\overset{ind.hyp.}{\Leftrightarrow} \mathscr{M}_R, w \models \varphi$ and $\mathscr{M}_R, w \models \psi \Leftrightarrow \mathscr{M}_R, w \models \varphi \wedge \psi$;
- $\mathscr{M}, w \models_T \neg \varphi \Leftrightarrow \mathscr{M}, w \models_F \varphi \overset{ind.hyp.}{\Leftrightarrow} \mathscr{M}_R, w \not\models \varphi \Leftrightarrow \mathscr{M}_R, w \models \neg \varphi$;
- $\mathscr{M}, w \models_T B\varphi \Leftrightarrow (\forall w' \in \mathscr{B}) \mathscr{M}, w' \models_T \varphi \overset{def.,ind.hyp.}{\Leftrightarrow} \mathscr{M}_R, w \models B\varphi$;
- $\mathscr{M}, w \models_T L\varphi \Leftrightarrow (\forall w' \in \mathscr{W}(\mathscr{B})) \mathscr{M}, w' \models_T \varphi \overset{def.,ind.hyp.}{\Leftrightarrow} \mathscr{M}_R, w \models L\varphi$. ■

Note that Levesque's use of a partial semantics cannot completely be captured by dispensing with a recursive truth definition in non-normal worlds. Supporting the truth at a situation may differ from truth at a (normal or non-normal) world: in general, $\mathscr{M}, w \models_T \varphi \not\Leftrightarrow \mathscr{M}_R, w \models \varphi$ ($\not\Rightarrow$: consider $w \in T(p) \cap F(p)$, $\varphi = p$; $\not\Leftarrow$: consider $w \in W^*$, $w \in T(p) \cap F(q)$, $V(p \wedge q, w) = 1$, $\varphi = p \wedge q$).

4.2. Fagin and Halpern's logic of awareness

The logic of awareness of Fagin and Halpern likewise is a logic of implicit and explicit belief, but now for the multi-agent case. Moreover, in its language, say $\mathbf{L}_2$, arbitrary nestings of intensional operators $L_1, \ldots, L_n$, $B_1, \ldots, B_n$ are permitted.[5] We now have *Kripke structures for awareness* $\mathscr{M} = \langle S, \pi, \mathscr{A}_1, \ldots, \mathscr{A}_n, \mathscr{B}_1, \ldots, \mathscr{B}_n \rangle$, where S is a (non-empty) set of states, π is a mapping from $S \times PRIM$ into $\{0, 1\}$, $\mathscr{B}_i (1 \leqslant i \leqslant n)$ is a serial, transitive, Euclidean relation on S,[6] and $\mathscr{A}_i (1 \leqslant i \leqslant n)$ is a mapping from S into 2^{PRIM}, assigning to each state s the set of those sentence letters of which agent i is *aware* at s. Besides a relation $\models$, Fagin and Halpern also define support relations $\models_T^\Psi$, $\models_F^\Psi$, where $\Psi \subseteq PRIM$:

- $\mathscr{M}, s \models_T^\Psi p \Leftrightarrow \pi(s, p) = 1$ and $p \in \Psi$,
  $\mathscr{M}, s \models_F^\Psi p \Leftrightarrow \pi(s, p) = 0$ and $p \in \Psi$,
  $\mathscr{M}, s \models p \Leftrightarrow \pi(s, p) = 1$;
- $\mathscr{M}, s \models_T^\Psi \varphi \wedge \psi \Leftrightarrow \mathscr{M}, s \models_T^\Psi \varphi$ and $\mathscr{M}, s \models_T^\Psi \psi$,
  $\mathscr{M}, s \models_F^\Psi \varphi \wedge \psi \Leftrightarrow \mathscr{M}, s \models_F^\Psi \varphi$ or $\mathscr{M}, s \models_F^\Psi \psi$,
  $\mathscr{M}, s \models \varphi \wedge \psi \Leftrightarrow \mathscr{M}, s \models \varphi$ and $\mathscr{M}, s \models \psi$;
- $\mathscr{M}, s \models_T^\Psi \neg \varphi \Leftrightarrow \mathscr{M}, s \models_F^\Psi \varphi$,
  $\mathscr{M}, s \models_F^\Psi \neg \varphi \Leftrightarrow \mathscr{M}, s \models_T^\Psi \varphi$,
  $\mathscr{M}, s \models \neg \varphi \Leftrightarrow \mathscr{M}, s \not\models \varphi$;

[5] Fagin and Halpern also include a constant *true*.

[6] In the case of veridical belief, $\mathscr{B}_i$ is required to be an equivalence relation on S.

- $\mathcal{M}, s \models_T^{\Psi} B_i \varphi \Leftrightarrow (\forall t \in S) \mathcal{B}_i st \Rightarrow \mathcal{M}, t \models_T^{\Psi \cap \mathcal{A}_i(s)} \varphi$,
  $\mathcal{M}, s \models_F^{\Psi} B_i \varphi \Leftrightarrow (\exists t \in s) \mathcal{B}_i st \wedge \mathcal{M}, t \models_F^{\Psi \cap \mathcal{A}_i(s)} \varphi$,
  $\mathcal{M}, s \models B_i \varphi \Leftrightarrow \mathcal{M}, s \models_T^{PRIM} B_i \varphi$;
- $\mathcal{M}, s \models_T^{\Psi} L_i \varphi \Leftrightarrow (\forall t \in S) \mathcal{B}_i st \Rightarrow \mathcal{M}, t \models_T^{\Psi} \varphi$,
  $\mathcal{M}, s \models_F^{\Psi} L_i \varphi \Leftrightarrow (\exists t \in S) \mathcal{B}_i st \wedge \mathcal{M}, t \models_F^{\Psi} \varphi$,
  $\mathcal{M}, s \models L_i \varphi \Leftrightarrow (\forall t \in S) \mathcal{B}_i st \Rightarrow \mathcal{M}, t \models \varphi$.

$\mathbf{L}_2 \ni \varphi$ is *valid in* $\mathcal{M} = \langle S, \pi, \mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{B}_1, \ldots, \mathcal{B}_n \rangle$ ($\mathcal{M} \models \varphi$) iff $(\forall s \in S)$ $\mathcal{M}, s \models \varphi$; φ is *valid* simpliciter ($\models \varphi$) iff for every Kripke structure for awareness $\mathcal{M}$, $\mathcal{M} \models \varphi$.

For $\varphi \in \mathbf{L}_2$, $\mathcal{M} = \langle S, \pi, \mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{B}_1, \ldots, \mathcal{B}_n \rangle$, and $w' \in S$, the sets $T_{\mathcal{M},w'}(\varphi)$, $F_{\mathcal{M},w'}(\varphi)$ are specified by simultaneous induction on φ as follows:

- $\varphi \in PRIM$: $T_{\mathcal{M},w'}(\varphi) = F_{\mathcal{M},w'}(\varphi) = \{\varphi\}$;
- $\varphi = \psi \wedge \chi$: $T_{\mathcal{M},w'}(\psi \wedge \chi) = T_{\mathcal{M},w'}(\psi) \cup T_{\mathcal{M},w'}(\chi)$,

$$F_{\mathcal{M},w'}(\psi \wedge \chi) = \begin{cases} F_{\mathcal{M},w'}(\psi) & \text{if } \mathcal{M}, w' \not\models \psi \text{ and } \mathcal{M}, w' \models \chi, \\ F_{\mathcal{M},w'}(\chi) & \text{if } \mathcal{M}, w' \not\models \chi \text{ and } \mathcal{M}, w' \models \psi, \\ F_{\mathcal{M},w'}(\psi) \cup F_{\mathcal{M},w'}(\chi) & \text{if } \mathcal{M}, w' \not\models \psi \text{ and } \mathcal{M}, w' \not\models \chi, \\ \varnothing & \text{otherwise;} \end{cases}$$

- $\varphi = \neg \psi$: $T_{\mathcal{M},w'}(\neg \psi) = F_{\mathcal{M},w'}(\psi)$, $F_{\mathcal{M},w'}(\neg \psi) = T_{\mathcal{M},w'}(\psi)$;
- $\varphi = B_i \psi$: $T_{\mathcal{M},w'}(B_i \psi) = \bigcup \{T_{\mathcal{M},t}(\psi) | \mathcal{B}_i w' t\}$,

$$F_{\mathcal{M},w'}(B_i \psi) = \begin{cases} F_{\mathcal{M},t}(\psi) \text{ for} & \text{if there is such a } t \in S, \\ \text{some } t \in S \text{ such that} & \\ \mathcal{B}_i w' t \text{ and } \mathcal{M}, t \models_F^{\mathcal{A}_i(w)} \psi, & \\ \varnothing & \text{otherwise;} \end{cases}$$

- $\varphi = L_i \psi$: $T_{\mathcal{M},w'}(L_i \psi) = \bigcup \{T_{\mathcal{M},t}(\psi) | \mathcal{B}_i w' t\}$,

$$F_{\mathcal{M},w'}(L_i \psi) = \begin{cases} F_{\mathcal{M},t}(\psi) \text{ for} & \text{if there is such a } t \in S, \\ \text{some } t \in S \text{ such that} & \\ \mathcal{B}_i w' t \text{ and } \mathcal{M}, t \not\models \psi, & \\ \varnothing & \text{otherwise.} \end{cases}$$

LEMMA 1. (i) $\mathcal{M}, w' \models_T^{\mathcal{A}_i(w)} \varphi \Leftrightarrow \mathcal{M}, w' \models \varphi$ *and* $T_{\mathcal{M},w'}(\varphi) \subseteq \mathcal{A}_i(w)$;
(ii) $\mathcal{M}, w' \models_F^{\mathcal{A}_i(w)} \varphi \Leftrightarrow \mathcal{M}, w' \not\models \varphi$ *and* $F_{\mathcal{M},w'}(\varphi) \subseteq \mathcal{A}_i(w)$.

PROOF. By simultaneous induction on $\varphi \in \mathbf{L}_2$. If $\varphi \in PRIM$, the assertions hold by definition.

- (i) $\varphi = \psi \wedge \chi$: $\mathcal{M}, w' \models \psi \wedge \chi$ and $T_{\mathcal{M},w'}(\psi \wedge \chi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
  $\mathcal{M}, w' \models \psi$, $\mathcal{M}, w' \models \chi$, and $T_{\mathcal{M},w'}(\psi) \cup T_{\mathcal{M},w'}(\chi) \subseteq \mathcal{A}_i(w) \overset{ind.hyp.}{\Leftrightarrow}$
  $\mathcal{M}, w' \models_T^{\mathcal{A}_i(w)} \psi$ and $\mathcal{M}, w' \models_T^{\mathcal{A}_i(w)} \chi \Leftrightarrow \mathcal{M}, w' \models_T^{\mathcal{A}_i(w)} \psi \wedge \chi$.
  (ii) $\mathcal{M}, w' \not\models \psi \wedge \chi$ and $F_{\mathcal{M},w'}(\psi \wedge \chi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
  $(\mathcal{M}, w' \not\models \psi$ and $F_{\mathcal{M},w'}(\psi) \subseteq \mathcal{A}_i(w))$ or
  $(\mathcal{M}, w' \not\models \chi$ and $F_{\mathcal{M},w'}(\chi) \subseteq \mathcal{A}_i(w)) \overset{ind.hyp.}{\Leftrightarrow}$
  $\mathcal{M}, w' \models_F^{\mathcal{A}_i(w)} \psi$ or $\mathcal{M}, w' \models_F^{\mathcal{A}_i(w)} \chi \Leftrightarrow \mathcal{M}, w' \models_F^{\mathcal{A}_i(w)} \psi \wedge \chi$.

- $\varphi = \neg\psi$: (i) $\mathcal{M}, w' \models \neg\psi$ and $T_{\mathcal{M},w'}(\neg\psi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
$\mathcal{M}, w' \not\models \psi$ and $F_{\mathcal{M},w'}(\psi) \subseteq \mathcal{A}_i(w) \overset{ind.hyp.}{\Leftrightarrow}$
$\mathcal{M}, w' \models_F^{\mathcal{A}_i(w)} \psi \Leftrightarrow \mathcal{M}, w' \models_T^{\mathcal{A}_i(w)} \neg\psi$.
(ii) Analogously.

- $\varphi = B_i\psi$: (i) $\mathcal{M}, w' \models B_i\psi$ and $T_{\mathcal{M},w'}(B_i\psi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
$\mathcal{M}, w' \models_T^{PRIM} B_i\psi$ and $T_{\mathcal{M},w'}(B_i\psi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
$(\forall t \in S)\big(\mathcal{B}_i w' t \Rightarrow \big(\mathcal{M}, t \models_T^{\mathcal{A}_i(w')} \psi$ and $T_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w)\big)\big) \overset{ind.hyp.}{\Leftrightarrow}$
$(\forall t \in S)\big(\mathcal{B}_i w' t \Rightarrow \big(\mathcal{M}, t \models \psi$ and $T_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w')$ and
$T_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w)\big)\big) \Leftrightarrow$
$(\forall t \in S)\big(\mathcal{B}_i w' t \Rightarrow \big(\mathcal{M}, t \models \psi$ and $T_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w') \cap \mathcal{A}_i(w)\big)\big) \overset{ind.hyp.}{\Leftrightarrow}$
$(\forall t \in S)(\mathcal{B}_i w' t \Rightarrow \mathcal{M}, t \models_T^{\mathcal{A}_i(w) \cap \mathcal{A}_i(w')} \psi) \Leftrightarrow \mathcal{M}, w' \models_T^{\mathcal{A}_i(w)} B_i\psi$;
(ii) $\mathcal{M}, w' \not\models B_i\psi$ and $F_{\mathcal{M},w'}(B_i\psi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
$\mathcal{M}, w' \not\models_T^{PRIM} B_i\psi$ and $F_{\mathcal{M},w'}(B_i\psi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
$(\exists t \in S)\big(\mathcal{B}_i w' t$ and $\mathcal{M}, t \models_F^{\mathcal{A}_i(w')} \psi$ and $F_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w)\big) \overset{ind.hyp.}{\Leftrightarrow}$
$(\exists t \in S)\big(\mathcal{B}_i w' t$ and $\mathcal{M}, t \not\models \psi$ and $F_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w)$ and
$F_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w')\big) \Leftrightarrow$
$(\exists t \in S)\big(\mathcal{B}_i w' t$ and $\mathcal{M}, t \not\models \psi$ and $F_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w) \cap \mathcal{A}_i(w')\big) \overset{ind.hyp.}{\Leftrightarrow}$
$(\exists t \in S)(\mathcal{B}_i w' t$ and $\mathcal{M}, t \models_F^{\mathcal{A}_i(w) \cap \mathcal{A}_i(w')} \psi) \Leftrightarrow \mathcal{M}, w' \models_F^{\mathcal{A}_i(w)} B_i\psi$;

- $\varphi = L_i\psi$: (i) $\mathcal{M}, w' \models L_i\psi$ and $T_{\mathcal{M},w'}(L_i\psi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
$(\forall t \in S)\big(\mathcal{B}_i w' t \Rightarrow \big(\mathcal{M}, t \models \psi$ and $T_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w)\big)\big) \overset{ind.hyp.}{\Leftrightarrow}$
$(\forall t \in S)(\mathcal{B}_i w' t \Rightarrow \mathcal{M}, t \models_T^{\mathcal{A}_i(w)} \psi) \Leftrightarrow \mathcal{M}, t \models_T^{\mathcal{A}_i(w)} L_i\psi$;
(ii) $\mathcal{M}, w' \not\models L_i\psi$ and $F_{\mathcal{M},w'}(L_i\psi) \subseteq \mathcal{A}_i(w) \Leftrightarrow$
$(\exists t \in S)\big(\mathcal{B}_i w' t$ and $\mathcal{M}, t \not\models \psi$ and $F_{\mathcal{M},t}(\psi) \subseteq \mathcal{A}_i(w)\big) \overset{ind.hyp.}{\Leftrightarrow}$
$(\exists t \in S)(\mathcal{B}_i w' t$ and $\mathcal{M}, t \models_F^{\mathcal{A}_i(w)} \psi) \Leftrightarrow \mathcal{M}, w' \models_F^{\mathcal{A}_i(w)} L_i\psi$. ■

CLAIM 2. *Every Kripke structure for awareness induces a Rantala model validating precisely the same* $\mathbf{L}_2$*-formulas.*

PROOF. Given a Kripke structure for awareness $\mathcal{M} = \langle S, \pi, \mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{B}_1, \ldots, \mathcal{B}_n\rangle$, define a Rantala model $\mathcal{M}_R = \langle W, W^*, R_{B_1}, \ldots, R_{B_n}, R_{L_1}, \ldots, R_{L_n}, V\rangle$ as follows:

- $W =_{def} S$;
- $W^* =_{def} \{\mathcal{A}_i(w) \mid 1 \leqslant i \leqslant n,\ w \in S\}$;
- $R_{B_i} =_{def} \mathcal{B}_i \cup \{\langle w, \mathcal{A}_i(w)\rangle \mid w \in S\}$, $1 \leqslant i \leqslant n$;
- $R_{L_i} =_{def} R_{B_i} \restriction W$;
- V: $FORM_{\mathbf{L}_2} \times (W \cup W^*) \rightarrow \{0, 1\}$ such that:
$\forall w \in W$,
$V(p, w) = 1 \Leftrightarrow \pi(w, p) = 1$,
$V(\varphi \wedge \psi, w) = 1 \Leftrightarrow V(\varphi, w) = V(\psi, w) = 1$,
$V(\neg\varphi, w) = 1 \Leftrightarrow V(\varphi, w) = 0$,
$V(B_i\varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)\big(R_{B_i} ww' \Rightarrow V(\varphi, w') = 1\big)$,
$V(L_i\varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)\big(R_{L_i} ww' \Rightarrow V(\varphi, w') = 1\big)$;
$\forall \mathcal{A}_i(w) \in W^*$,
$V\big(\varphi, \mathcal{A}_i(w)\big) = 1 \Leftrightarrow \bigcup \{T_{\mathcal{M},w'}(\varphi) \mid \mathcal{B}_i ww'\} \subseteq \mathcal{A}_i(w)$.

We show by induction on $\varphi \in \mathbf{L}_2$ that $(\forall \varphi \in \mathbf{L}_2)\ \mathcal{M} \models \varphi \Leftrightarrow \mathcal{M}_R \models \varphi$, whence the claim follows. Suppose that $w \in S$; the only interesting case is $\varphi = B_i \psi$.

$$\mathcal{M}, w \models B_i \psi \Leftrightarrow \mathcal{M}, w \models_T^{PRIM} B_i \psi \Leftrightarrow$$
$$(\forall w' \in S)(\mathcal{B}_i ww' \Rightarrow \mathcal{M}, w' \models_T^{\mathcal{A}_i(w)} \psi) \overset{Lemma\,1}{\Leftrightarrow}$$
$$(\forall w' \in S)\big(\mathcal{B}_i ww' \Rightarrow \big(\mathcal{M}, w' \models \psi \text{ and } T_{\mathcal{M},w'}(\psi) \subseteq \mathcal{A}_i(w)\big)\big) \overset{ind.hyp.}{\Leftrightarrow}$$
$$(\forall w' \in S)\big(\mathcal{B}_i ww' \Rightarrow \big(\mathcal{M}_R, w' \models \psi \text{ and } V(\psi, \mathcal{A}_i(w)) = 1\big)\big) \Leftrightarrow \mathcal{M}_R, w \models B_i \psi. \quad \blacksquare$$

4.3. Fagin and Halpern's logic of general awareness

In Fagin and Halpern's logic of general awareness, a richer language, say $\mathbf{L}_3$, is used, extending the vocabulary of $\mathbf{L}_2$ by unary awareness operators $A_1, \ldots, A_n$. A *Kripke structure for general awareness* is a structure $\mathcal{M} = \langle S, \pi, \mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{B}_1, \ldots, \mathcal{B}_n \rangle$, where again S is a (non-empty) set of states, π is a mapping from $S \times PRIM$ into $\{0, 1\}$, and $\mathcal{B}_i (1 \leqslant i \leqslant n)$ is a serial, transitive, Euclidean (or an equivalence) relation on S. $\mathcal{A}_i (1 \leqslant i \leqslant n)$, however, now maps states s into sets of $\mathbf{L}_3$-formulas. The truth definition looks like this:

- $\mathcal{M}, s \models p \Leftrightarrow \pi(s, p) = 1$,
- $\mathcal{M}, s \models \varphi \wedge \psi \Leftrightarrow \mathcal{M}, s \models \varphi$ and $\mathcal{M}, s \models \psi$,
- $\mathcal{M}, s \models \neg \varphi \Leftrightarrow \mathcal{M}, s \not\models \varphi$,
- $\mathcal{M}, s \models B_i \varphi \Leftrightarrow \varphi \in \mathcal{A}_i(s)$ and $(\forall t \in S) \mathcal{B}_i st \Rightarrow \mathcal{M}, t \models \varphi$,
- $\mathcal{M}, s \models L_i \varphi \Leftrightarrow (\forall t \in S) \mathcal{B}_i st \Rightarrow \mathcal{M}, t \models \varphi$,
- $\mathcal{M}, s \models A_i \varphi \Leftrightarrow \varphi \in \mathcal{A}_i(s)$.

CLAIM 3. *Every Kripke structure for general awareness induces a Rantala model validating precisely the same $\mathbf{L}_3$-formulas.*

PROOF. Given a Kripke structure for general awareness $\mathcal{M} = \langle S, \pi, \mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{B}_1, \ldots, \mathcal{B}_n \rangle$, define a Rantala model $\mathcal{M}_R = \langle W, W^*, R_{B_1}, \ldots, R_{B_n}, R_{L_1}, \ldots, R_{L_n}, V \rangle$ as follows:

- $W =_{def} S$;
- $W^* =_{def} \{\mathcal{A}_i(w) \mid 1 \leqslant i \leqslant n, w \in S\}$;
- $R_{B_i} =_{def} \mathcal{B}_i \cup \{\langle w, \mathcal{A}_i(w) \rangle \mid w \in S\}$, $1 \leqslant i \leqslant n$;
- $R_{L_i} =_{def} R_{B_i} \restriction W$;
- V: $FORM_{\mathbf{L}_3} \times (W \cup W^*) \rightarrow \{0, 1\}$ such that:
  $\forall w \in W$,
  $V(p, w) = 1 \Leftrightarrow \pi(w, p) = 1$,
  $V(\varphi \wedge \psi, w) = 1 \Leftrightarrow V(\varphi, w) = V(\psi, w) = 1$,
  $V(\neg \varphi, w) = 1 \Leftrightarrow V(\varphi, w) = 0$,
  $V(B_i \varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)\big(R_{B_i} ww' \Rightarrow V(\varphi, w') = 1\big)$,
  $V(L_i \varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)\big(R_{L_i} ww' \Rightarrow V(\varphi, w') = 1\big)$,
  $V(A_i, w) = 1 \Leftrightarrow \varphi \in \mathcal{A}_i(w)$;

$\forall \mathscr{A}_i(w) \in W^*$,
$V(\varphi, \mathscr{A}_i(w)) = 1 \Leftrightarrow \varphi \in \mathscr{A}_i(w)$.

Obviously, $(\forall \varphi \in \mathbf{L}_3)\, \mathscr{M} \models \varphi \Leftrightarrow \mathscr{M}_R \models \varphi$. ■[7]

4.4. Fagin and Halpern's logic of local reasoning

Pursuing the idea that agents may be conceived of as 'societies of minds', Fagin and Halpern present still another logic of implicit and explicit belief. This logic is characterised by *Kripke structures for local reasoning* $\mathscr{M} = \langle S, \pi, \mathscr{C}_1, \ldots, \mathscr{C}_n \rangle$, where again S is a (non-empty) set of states and π a mapping from $S \times PRIM$ into $\{0, 1\}$. $\mathscr{C}_i (1 \leqslant i \leqslant n)$ is a mapping from S into 2^{2^S}. $\mathscr{C}_i(s)$ can be viewed as comprising the different sets of states agent i considers possible at s, the sets forming her or his 'society of minds' at s. The truth relation $\models$ is inductively defined as follows:

- $\mathscr{M}, s \models p \Leftrightarrow \pi(s, p) = 1$,
- $\mathscr{M}, s \models \varphi \wedge \psi \Leftrightarrow \mathscr{M}, s \models \varphi$ and $\mathscr{M}, s \models \psi$,
- $\mathscr{M}, s \models \neg \varphi \Leftrightarrow \mathscr{M}, s \not\models \varphi$,
- $\mathscr{M}, s \models B_i \varphi \Leftrightarrow (\exists X \in \mathscr{C}_i(s))((\forall t \in X)\, \mathscr{M}, t \models \varphi)$,
- $\mathscr{M}, s \models L_i \varphi \Leftrightarrow (\forall t \in \bigcap_{X \in \mathscr{C}_i(s)} X)\, \mathscr{M}, t \models \varphi$.[8]

CLAIM 4. *Every Kripke structure for local reasoning induces a Rantala model validating precisely the same* $\mathbf{L}_2$*-formulas.*

PROOF. Given a Kripke structure for local reasoning $\mathscr{M} = \langle S, \pi, \mathscr{C}_1, \ldots, \mathscr{C}_n \rangle$, define a Rantala model $\mathscr{M}_R = \langle W, W^*, R_{B_1}, \ldots, R_{B_n}, R_{L_1}, \ldots, R_{L_n}, V \rangle$ as follows:

- $W =_{def} S$;
- $W^* =_{def} \{w_i | w \in W, 1 \leqslant i \leqslant n\}$, where the w_i's are new reference points;
- $R_{B_i} =_{def} \{\langle w, w' \rangle | w \in W, w' \in \bigcap_{X \in \mathscr{C}_i(w)} X\} \cup \{\langle w, w_i \rangle | w \in W\}$;
- $R_{L_i} =_{def} R_{B_i} \upharpoonright W$;
- V: $FORM_{\mathbf{L}_2} \times (W \cup W^*) \to \{0, 1\}$ such that:
  $\forall w \in W$,
  $V(p, w) = 1 \Leftrightarrow \pi(w, p) = 1$,
  $V(\neg \varphi, w) = 1 \Leftrightarrow V(\varphi, w) = 0$,
  $V(\varphi \wedge \psi, w) = 1 \Leftrightarrow V(\varphi, w) = V(\psi, w) = 1$,

[7] As Joe Halpern (private communication) observed, in many cases a Rantala model gives rise to an equivalent Kripke structure for general awareness: "given a Rantala model, take a Kripke structure consisting of precisely the normal worlds. Then define the awareness set at a world w to consists precisely of those formulas φ such that $K(\varphi)$ is true at that state in the Rantala model". However, in general this construction doesn't work: consider e.g. the canonical Rantala model for non-modal classical propositional logic.

[8] In order to model veridical belief, it has to be assumed that $(\forall s \in S)(\forall X \in \mathscr{C}_i(s))\, s \in X$.

$V(B_i\varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)(R_{B_i}ww' \Rightarrow V(\varphi, w') = 1)$,
$V(L_i\varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*)(R_{L_i}ww' \Rightarrow V(\varphi, w') = 1)$;
$\forall w_i \in W^*$,
$V(\varphi, w_i) = 1 \Leftrightarrow \mathcal{M}, w \models B_i\varphi$.

We show that $(\forall\varphi \in \mathbf{L}_2)\mathcal{M} \models \varphi \Leftrightarrow \mathcal{M}_R \models \varphi$. Assume that $w \in W$. Consider $\varphi = B_i\psi$. $M_R, w \models \varphi \Leftrightarrow \mathcal{M}_R, w \models L_i\psi$ and $\mathcal{M}, w \models B_i\psi$. But $\mathcal{M} \models B_i\psi \rightarrow L_i\psi$. ■

Fagin and Halpern (FH, p. 69) point out that explicit belief in their logic of local reasoning precisely satisfies the axioms of (the multi-modal version of) the classical modal propositional logic *EMNP* (cf. [Ch]). The idea to take the minimal classical modal logic *E* as a kind of epistemic base logic can be found in [Va], where Vardi moreover presents a *constructive* model theory for (extensions of) *E* in terms of what he calls *belief worlds*. Vardi shows (Theorem 5) that, given the semantics of belief worlds, one may construct a multi-modal neighbourhood model $\mathcal{M}_b = \langle W_\omega, N, \Pi\rangle$ such that for every $\varphi \in \mathbf{L}$, φ is true at world $f \in W_\omega$ iff $f \models \varphi$, where W_ω is the set of infinitary belief worlds. In [PW, p. 13] it has been observed that every neighbourhood model induces an equivalent Rantala model, hence so does $\mathcal{M}_b$. Therefore, given the semantics of belief worlds, one may construct a canonical Rantala model for *E*.

4.5. Van der Hoek and Meyer's logic of awareness and principles

Van der Hoek and Meyer's logic of awareness and principles (or 'prejudices') ([vdHM]) is an extension of Fagin and Halpern's logic of general awareness. The idea is to consider in addition something like 'duals' $\mathcal{P}_1, \ldots, \mathcal{P}_n$ of the awareness functions $\mathcal{A}_1, \ldots, \mathcal{A}_n$. The principle functions $\mathcal{P}_1, \ldots, \mathcal{P}_n$ are used to model 'reasoning against the facts'. Syntactically, they are reflected by principle operators $P_1, \ldots, P_n$, providing a language $\mathbf{L}_4$. Thus, one obtains *Kripke structures for awareness and principles* $\mathcal{M} = \langle S, \pi, \mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{P}_1, \ldots, \mathcal{P}_n, \mathcal{B}_1, \ldots, \mathcal{B}_n\rangle$, where, as before, S is a (non-empty) set of states, π is a mapping from $S \times PRIM$ into $\{0, 1\}$, and $\mathcal{B}_i (1 \leqslant i \leqslant n)$ is a serial, transitive, Euclidean (or an equivalence) relation on S. The $\mathcal{A}_i, \mathcal{P}_i (1 \leqslant i \leqslant n)$ map states s into sets of $\mathbf{L}_4$-formulas. $P_i\varphi$ is true in $\mathcal{M}$ at $s \in S$ ($\mathcal{M}, s \models P_i\varphi$) iff $\varphi \in \mathcal{P}_i(s)$. The clause for $L_i\varphi$ now becomes

- $\mathcal{M}, s \models L_i\varphi \Leftrightarrow \varphi \in \mathcal{P}_i(s)$ or $(\forall t \in S)\mathcal{B}_i st \Rightarrow \mathcal{M}, t \models \varphi$.

CLAIM 5. *Every Kripke structure for awareness and principles induces a Rantala model validating precisely the same* $\mathbf{L}_4$*-formulas.*

PROOF. Some straightforward adjustments are required of the earlier construction of a Rantala model, given a Kripke model for general awareness. We now define W^* as $W_A \cup W_P$, where $W_A = \{\mathcal{A}_i(w) \mid 1 \leqslant i \leqslant n, w \in S\}$, $W_P = \{\mathcal{P}_i(w) \mid 1 \leqslant i \leqslant n, w \in S\}$. $R_{B_i} =_{def} \mathcal{B}_i \cup \{\langle w, \mathcal{A}_i(w)\rangle \mid w \in S\}$; we postulate that $R_{L_i} = \mathscr{B}_i$ or $R_{L_i} = \{\langle w, \mathscr{P}_i(w)\rangle | w \in S\}$. In normal worlds w, formulas $B_i\varphi$, $L_i\varphi$, $A_i\varphi$, $P_i\varphi$ are evaluated as follows:

- $V(B_i\varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*) R_{B_i} ww' \Rightarrow V(\varphi, w) = 1$,
- $V(L_i\varphi, w) = 1 \Leftrightarrow (\forall w' \in W \cup W^*) R_{L_i} ww' \Rightarrow V(\varphi, w) = 1$,
- $V(A_i\varphi, w) = 1 \Leftrightarrow \varphi \in \mathscr{A}_i(w)$,
- $V(P_i\varphi, w) = 1 \Leftrightarrow \varphi \in \mathscr{P}_i(w)$.

For non-normal worlds $\mathscr{A}_i(w) \in W_A$, $\mathscr{P}_i(w) \in W_P$, we assume that $V(\varphi, \mathscr{A}_i(w)) = 1 \Leftrightarrow \varphi \in \mathscr{A}_i(w)$, $V(\varphi, \mathscr{P}_i(w)) = 1 \Leftrightarrow \varphi \in \mathscr{P}_i(w)$. It is then easily seen that $(\forall \varphi \in \mathbf{L}_4)\mathscr{M} \models \varphi \Leftrightarrow \mathscr{M}_R \models \varphi$. ■

5. Discussion

So far, non-normal worlds semantics has been *motivated* as a basic and general possible worlds framework for reasoning about knowledge and belief. Moreover, it has been shown that the model theories of various logics of knowledge and belief which successfully avoid particular aspects of logical omniscience in fact all fit neatly into the framework of non-normal worlds.[9] Although the non-normal worlds framework provides a common perspective behind a variety of epistemic logics, there still is a number of criticisms and reservations concerning non-normal worlds semantics that deserve some comment.

What actually *are* non-normal worlds? Or at least, how can one make sense of them? Non-normal worlds *are* theoretical entities used e.g. in epistemic logic; they are reference points at which truth need not be recursively specified. If this 'definitiorial' answer is unsatisfactory, it might, I suppose, in the first place be due to the kind of question and not to this particular answer. A more reasonable question is, how to make sense of non-normal worlds. The epistemic interpretation of necessity reads like this: "agent i knows at world s that φ iff φ is true at all of i's epistemic alternatives from s (true at all worlds 'compatible' with what i knows at s)".[10] There is no reason to assume that epistemic alternatives necessarily display some logical closure properties. On the contrary, one should expect epistemic alternatives that are not models of e.g. classical propositional logic. Moreover, there is nothing obscure about

---

[9] This may also prove useful for more specific research problems. As pointed out in [Ko] and to some extent investigated in [vdHM], questions of first-order definability of modal formulas become far more complicated and difficult to survey, if awareness (and principle) functions are present. Non-normal worlds semantics may be expected to provide a still involved, yet uniform perspective on first-order definability by means of a (highly non-standard) *correspondence theory* (cf. [vB], [PW]).

[10] Unfortunately, it seems that the idea of 'compatibility' has never been explicitly worked out.

assuming 'inconsistent' epistemic alternatives, verifying some formula and its negation. Inconsistent epistemic alternatives need, of course, not be known or believed to be inconsistent; allowing for such worlds (knowledge states) enables one e.g. to interpret epistemic alternatives as possibly inconsistent data bases. A genuine option here is going partial. Should one assume an epistemic alternative to be *complete* in the sense of deciding every (primitive) sentence? If not, there is still the question of a recursive truth definition. It was doing without a recursive definition of truth at non-normal worlds what allowed us to treat certain attitude verbs as descriptive expressions. A partial logic of knowledge or belief, after all, remains a *logic* of knowledge or belief. It has e.g. been criticised (cf. [FH], [Va]) that an agent in Levesque's logic of implicit and explicit belief still is a perfect reasoner wrt the relevance logic of first degree entailments (cf. e.g. [Du]) in so far as φ entails ψ iff explicit belief in φ implies explicit belief in ψ.

As expected, the reservations against non-normal worlds semantics to be found in the literature centre around the intuitions about non-normal worlds. In his survey [Ha1, p. 7] e.g., Halpern regrets that "impossible worlds have not been very well motivated". Yet, commenting on non-normal worlds semantics and syntactic models of knowledge and belief, Fagin and Halpern are somewhat undecided:

> The syntactic approach lacks the elegance and *intuitive appeal* of the semantic approach. However, the semantic rules used to assign truth values to the logical connectives in the impossible-worlds approach have tended to be *nonintuitive*, and it is not clear to what extent this approach has been successful in truly capturing our intuitions about knowledge and belief. ([FH, p. 40], my emphasis)

Whatever our intuitions about knowledge and belief are, it has been shown above that the varieties of knowledge and belief discussed in [FH] *can* truly be captured by non-normal worlds semantics. Fagin and Halpern's article may therefore be read as providing particular versions of non-normal worlds semantics. Thus, in Rantala models where for each normal world w there is precisely one non-normal world to which w is R_{B_i}-related, non-normal worlds may be interpreted as awareness sets. The restrictions on awareness functions A_i in [FH] then directly translate into constraints on evaluation in non-normal worlds.

Finally it should be mentioned that non-normal worlds are by no means bound to be interpreted syntactically. Rantala ([Ra1], [Ra2]) e.g. briefly suggests considering non-normal worlds structures $\langle W, W^*, R_1, \ldots, R_n, V\rangle$, where W^* is a set of models of intuitionistic (first order) logic such that for every $w^* \in W^*$ and every non-modal φ, $V(\varphi, w^*) = 1$ iff $w^* \models \varphi$ intuitionistically. In general, however, the Rantala semantics constitutes an *explanatory* framework for reasoning about knowledge and belief, simply because it 'realistically' allows for epistemic alternatives other than normal possible worlds.

Acknowledgement

I gratefully acknowledge support obtained from the Senat von Berlin and from the Studienstiftung des deutschen Volkes. Moreover, I would like to thank Johan van Benthem, David Pearce, Yde Venema, Gerd Wagner, and, in particular, Wiebe van der Hoek and an anonymous referee for various remarks on earlier drafts of this paper.

References

[Ba] J. BARWISE, *The Situation in Logic — IV: On the Model Theory of Common Knowledge*, Center for the Study of Language and Information, Report No. CSLI-88-122, Stanford, 1988.

[vB] J. F. A. K. VAN BENTHEM, *Correspondence theory*, in [GG], vol. II, 1984, pp. 167–247.

[Ca] R. CARNAP, ***Meaning and Necessity***, Chicago University Press, 1947.

[Ch] B. F. CHELLAS, ***Modal Logic: An Introduction***, Cambridge University Press, 1980.

[Cr1] M. J. CRESSWELL, *Intensional logics and logical truth*, ***Journal of Philosophical Logic*** 1 (1972), pp. 2–15.

[Cr2] M. J. CRESSWELL, *Hyperintensional logic*, ***Studia Logica*** 34 (1975), pp. 25–38.

[Du] J. M. DUNN, *Relevance logic and entailment*, in [GG], vol. III, 1986, pp. 117–224.

[FH] R. FAGIN and J. Y. HALPERN, *Belief, awareness, and limited reasoning*, ***Artificial Intelligence*** 34 (1988), pp. 39–76.

[GG] D. GABBAY and F. GUENTNER (eds.), ***Handbook of Philosophical Logic***, vols. I–IV, Reidel, Dordrecht, 1984–1989.

[Ha1] J. Y. HALPERN, *Reasoning about knowledge: An overview*, in [Ha2], pp. 1–17.

[Ha2] J. Y. HALPERN (ed.), ***Theoretical Aspects of Reasoning about Knowledge: Proceedings of the 1986 Conference***, Morgan Kaufmann, Los Altos, 1986.

[Hi1] J. HINTIKKA, ***Knowledge and Belief: An Introduction to the Logic of the Two Notions***, Cornell University Press, 1962.

[Hi2] J. HINTIKKA, *Impossible possible worlds vindicated*, ***Journal of Philosophical Logic*** 4 (1975), pp. 475–484.

[vdHM] W. VAN DER HOEK and J.-J. Ch. MEYER, *Possible logics for belief*, Rapport IR-170, Vrije Universiteit Amsterdam, 1988.

[Ko] K. KONOLIGE, *What awareness isn't: A sentential view of implicit and explicit belief*, in [Ha2], pp. 241–250.

[Le] H. J. LEVESQUE, *A logic of implicit and explicit belief*, ***Proceedings AAAI-84***, Austin, Texas, 1984, pp. 198–202.

[NS] I. NIINILUOTO and E. SAARINEN (eds.), ***Intensional Logic: Theory and Applications (Acta Philosophica Fennica*** 35), Helsinki, 1982.

[Pa] R. PARIKH, *Knowledge and the problem of logical omniscience*, in [RZ], pp. 432–439.

[PW] D. PEARCE and H. WANSING, *On the methodology of possible worlds semantics, II: Non-normal worlds and propositional attitudes*, Berichte der Gruppe für Logik, Wissenstheorie und Information, Freie Universität Berlin, Bericht 2/89, 1989, forthcoming in ***Acta Philosophica Fennica.***

[Ra1] V. RANTALA, *Impossible worlds semantics and logical omniscience*, in [NS], pp. 106–115.

[Ra2] V. RANTALA, *Quantified modal logic: Non-normal worlds and propositional attitudes*, ***Studia Logica*** 41 (1982), pp. 41–65.

[RZ] Z. W. RAS and M. ZEMANKOVA (eds.), ***Proceedings of the AMC SIGART International Symposium on Methodologies for Intelligent Systems***, Knoxville, Tennessee, 1986.

[Sa] E. SAARINEN, *Propositional attitudes are not attitudes towards propositions*, in [NS], pp. 130–162.
[Va] M. Y. VARDI, *On epistemic logic and logical omniscience*, in [Ha2], pp. 293–305.
[Wa] H. WANSING, *Bemerkungen zur Semantik nicht-normaler möglicher Welten*, ***Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*** 35 (1989), pp. 551–557.

GRUPPE FÜR LOGIK,
WISSENSTHEORIE UND INFORMATION
FU BERLIN
HABELSCHWERDTER ALLEE 30
D 1000 BERLIN 33
GERMANY

Received November 15, 1989
Revised March 23, 1990