Epistemic Logics of Argument

by Matthew P. Wampler-Doty

Abstract

In [11], Johan van Benthem proposed a research program for logic of *explicit knowledge*. Two efforts have since emerged with proposals for logic of explicit knowledge: the *Dynamics of Awareness* [18] and *Justification Logic* [1]. The purpose of this paper is to propose novel, argument based semantics for various modal logics related to these previous efforts, for use in reasoning about knowledge bases. We conclude with an application to naturalized epistemology. The results in this paper are based on previous work carried out in (FIXME).

1 Introduction

The modern tradition in epistemic logic is to assume knowledge modalities conform to the $S5$ axiom schema. As noted in [10, 15], the semantics of $S5$ knowledge correspond exactly to partitioning a collection of situations into *information sets*, which is the tradition in game theory and decision theory. While it is not commonly acknowledged in epistemic logic, economists and philosophers accept that traditional decision theory is externalist and behaviourist in nature[1].

In [11], Johan van Benthem proposed a research program to find logics for explicit knowledge, providing the first suggestion of an internalist perspective on epistemic logic. Subsequently, Sergei Artemov and Elena Nogina proposed a logic of explicit justification, which has come to be known as *Justification Logic* (JL) [1], based on Artemov's *Logic of Proofs* [2]. While the original semantics of JL was interpretability into Peano Arithmetic, Melvin Fitting proposed Kripke semantics for JL in [9]. Recently, Fernando Velázquez-Quesada and Johan van Benthem have developed a simpler framework for explicit epistemics in [18], entitled the *Dynamics of Awareness*. This work was based on Joseph Halpern and Ronald Fagin's original *Awareness Logic* [8].

In this essay we repurpose various modal logics to take on an internalist reading. The concept of a *knowledge base*, from which beliefs may be implicitly deduced, will play a crucial role in our discussion. We propose this as an avenue for representing foundationalist perspectives on epistemology in epistemic logic. Our philosophical motivation is taken from two sources. The first is Vincent Hendricks in [11], where he characterizes the principal of *logical omniscience* for implicit knowledge in epistemic logic[2]:

> *Whenever an agent Ξ knows all of the formulae in $\mathcal{A}$, and ϕ follows logically from $\mathcal{A}$, then Ξ [implicitly] knows ϕ.*

We will design our semantics such that "$\Box\phi$" may be equated with "ϕ follows logically from a knowledge base $\mathcal{A}$," which is sometimes written as $\phi \in \mathrm{Cn}(\mathcal{A})$ in the artificial intelligence literature. Our second inspiration comes from what Hilary Kornblith calls the *the arguments on paper thesis* [12], which he feels characterizes apsychological, internalist theories of knowledge:

> Let us suppose that, for any person, it is possible, at least in principle, to list all of the propositions that person believes. The arguments-on-paper thesis is just the view that a person has a justified belief that a particular proposition is true just in case that proposition appears on the list of propositions that person believes, and either it requires no argument, or a *good argument* can be given for it which takes as premises certain other propositions on the list.

Kornblith asserts that foundationalism and coherentism vie for accounts of a "good argument" in the above thesis. He provides an extensive bibliography citing proposals for this principle by a number of 20th century epistemologists, including figures such as A.J. Ayer and C.I. Lewis. The rest of Kornblith's paper is devoted to attacking this view and proposing a form of naturalized epistemology; we will not address this debate here, however.

1. An early essay by Amartya Sen on the philosophical foundations of traditional decision theory makes the behaviorist reading of decision theory clear [17]. Kaushik Basu also discusses the behaviorist nature of decision theory [3, pgs. 53–54]. Finally, Donald Davidson appeals to decision theory to motivate an externalist epistemology in [7].

2. We have modified Hendricks notation here slightly to match our own.

We adopt a arguments-on-paper-thesis perspective on epistemic logic in this paper. We consider a *good argument* to be a logical derivation from propositions present in a knowledge base. In this fashion, our Special attention will be given to *sound derivations*, which will be thought of as a form of knowledge. In Awareness Logic, we will interpret awareness of a formula as membership in a knowledge base. We prove completeness for basic awareness logic and a hybrid logic extension. Logics of multiple knowledge bases are also presented: a simplified form of Justification Logic, a logic with neighborhood semantics, an a logic with modalities for quantifying over knowledge bases. We conclude with an application to naturalized epistemology found in the psychological literature.

2 Basic Explicit Logics of Knowledge Bases

2.1 Awareness Logic

In this section, we give the explicit semantics to basic awareness logic, originally presented in [8], with an additional axiom. One difference in our presentation is the inclusion of a novel letter and a superficial change in notation. This work is inspired by the developments in van Benthem and Velázquez-Quesada [18].

Definition 1. *Let Φ be a set of letters and define the language $\mathcal{L}_A(\Phi)$ as:*

$$\phi ::= p \in \Phi \mid \circlearrowleft \mid \bot \mid \phi \to \psi \mid \Box\phi \mid A{:}\,\phi$$

We define the other connectives and modal operators as usual for classical modal logics.

Definition 2. *An* ***awareness model*** $\mathbb{M} = \langle W, V, R, \mathcal{A}\rangle$ *is a Kripke model with a valuation $V\colon \Phi \cup \{\circlearrowleft\} \to 2^W$ and an awareness function $\mathcal{A}\colon W \to 2^{\mathcal{L}_A(\Phi)}$. We will usually denote $\mathcal{A}(w)$ as $\mathcal{A}_w$.*

Define the semantic entailment relation $\vDash$ where atomic propositions, logical connectives and modality are interpreted as usual. We employ the following novel definition:

$$\mathbb{M}, w \vDash A{:}\,\phi \quad \textit{iff} \quad \phi \in \mathcal{A}_w$$

The following definition distinguishes us from previously investigations of awareness logic:

Definition 3. *Define the "**theory of a model** $\mathbb{M}$" as:*

$$\mathrm{Th}(\mathbb{M}) := \{\phi \mid \mathbb{M}, v \vDash \phi \textit{ for all } v \in W\}$$

We define the following properties that a model $\mathbb{M}$ may posses:

- ***CSQ*** — $\mathbb{M}, w \vDash \Box\phi$ *iff* $\mathrm{Th}(\mathbb{M}) \cup \mathcal{A}_w \vdash \phi$
- ***SND*** — $\mathbb{M}, w \vDash \circlearrowleft$ *iff* $\mathbb{M}, w \vDash \phi$ *for all* $\phi \in \mathcal{A}_w$

ere $\vdash$ is any sound logical consequence relation for $\vDash$ with ***modus ponens*** *and* ***reflection*** *(ie, $\Gamma \vdash \phi$ if $\phi \in \Gamma$).*

Intuitively, the **CSQ** asserts that an agent believes a formula ϕ if and only if it follows from their knowledge base and their background knowledge, represented by $\mathrm{Th}(\mathbb{M})$. **CSQ** is Vincent Hendrik's principle of logical omniscience mentioned in §1.

SND asserts that $\circlearrowleft$ corresponds to the agent's knowledge base being sound. A sound knowledge base will only render true conclusions when used in deductions. If one equates knowledge with "The existance of a sound deduction" then $\circlearrowleft$ is a mechanism for investigating this notion. That is, if $\mathbb{M}, w \vDash \circlearrowleft \wedge \Box\phi$, this is sufficient for the agent (implicitly) *knowing* ϕ on the basis of $\mathcal{A}_w$, rather than merely believing ϕ on the basis of $\mathcal{A}_w$.

Neither **CSQ** nor **SND** correspond to simple modally definable properties.

$$\begin{array}{l}
\vdash \phi \to \psi \to \phi \\
\vdash (\phi \to \psi \to \chi) \to (\phi \to \psi) \to \phi \to \chi \\
\vdash ((\phi \to \bot) \to (\psi \to \bot)) \to \psi \to \phi \\
\vdash \Box(\phi \to \psi) \to \Box\phi \to \Box\psi \\
\vdash A{:}\,\phi \to \Box\phi \\
\vdash \circlearrowleft \to \Box\phi \to \phi
\end{array}$$

$$\frac{\vdash \phi \to \psi \qquad \vdash \phi}{\vdash \psi} \qquad\qquad \frac{\vdash \phi}{\vdash \Box\phi}$$

Table 1. An awareness logic for $\mathbf{CSQ}_{\vdash}$ and **SND**

Theorem 4. *Assuming an infinite store of proposition letters Φ, the logic in Table 1 is sound and weakly complete for awareness models making true* ***CSQ*** *and* ***SND***

Proof. Soundness is straightforward, so we will only address completeness. Assume $\nvdash \psi$. Consider the finite canonical model $\mathbb{M} = \langle W, V, R, \mathcal{A} \rangle$ formed of maximally consistent sets of subformulae of ϕ (closed under pseudo-negation), as per the modal completeness proofs suggested in [5, chapter 5]. We have that $\mathbb{M}, w \nvDash \psi$ for some world w. Moreover, it is straightforward to verify that $\mathbb{M}$ makes true

1. W is finite and $\mathcal{A}_v$ is finite for all $v \in W$

2. if $\mathbb{M}, v \vDash A{:}\, \phi$ then $\mathbb{M}, v \vDash \Box \phi$

3. if $\mathbb{M}, v \vDash \circlearrowleft$ then $v R v$

We next produce a finite bisimular model $\mathbb{M}'$ which makes true (1), (2) and a stronger from of (3):

3′. $\mathbb{M}', v \vDash \circlearrowleft$ iff $v R v$

To this end define $\mathbb{M}' := \langle W', V', R', \mathcal{A}' \rangle$ such that

- $W' := W \uplus W$; furthermore, let l, r be the two canonical injections associated with the coproduct $W \uplus W$. We denote $l(v)$ and $r(v)$ as v_l and v_r respectively
- $V'(p) := \{v_l, v_r \mid v \in V(p)\}$
- $R' := \{(v_l, u_r), (v_r, u_l) \mid v R u\} \cup \{(v_l, v_l), (v_r, v_r) \mid \mathbb{M}, v \vDash \circlearrowleft\}$
- $\mathcal{A}'(v_l) := \mathcal{A}'(v_r) := \mathcal{A}(v)$

It is straightforward to verify that $\mathbb{M}'$ makes true the desired properties. If we let $Z := \{(v, v_l), (v, v_r) \mid v \in W\}$, then Z is a bisimulation between $\mathbb{M}$ and $\mathbb{M}'$. Therefore we know there is some $w \in W$ such that $\mathbb{M}', w_l \nvDash \psi$ and $\mathbb{M}', w_r \nvDash \psi$.

Finally we construct a model $\mathbb{M}''$ which agrees with $\mathbb{M}'$ for all subformulae of ψ, which makes true **CSQ** and **SND**. Let Λ be the set of proposition letters occuring in ψ. We will make use members of $\Phi \backslash \Lambda$ as *nominals*, as per the tradition in hybrid logic. This is possible since both W' and Λ are finite and Φ is infinite by hypothesis. Let $\iota{:}\, W \uplus W \hookrightarrow \Phi \backslash \Lambda$ be an injection assigning a fresh nominal associated with each world in $\mathbb{M}'$. Now define $\mathbb{M}'' := \langle W'', V'', R'', \mathcal{A}'' \rangle$, where:

- $W'' := W'$
- $V''(p) := \begin{cases} V'(p) & p \in \Lambda \\ \{v\} & p = \iota(v) \\ \varnothing & o/w \end{cases}$
- $R'' := R'$
- $\mathcal{A}''_v := \{\phi \mid \phi \in \mathcal{A}'_v \text{ and } \phi \text{ is a subformula of } \psi\} \cup \{\neg \iota(u) \mid \neg v R' u\}$

It is easy to verify that for every subformula ϕ of ψ that $\mathbb{M}', v \vDash \phi$ if and only if $\mathbb{M}'', v \vDash \phi$, hence $\mathbb{M}'', w_l \nvDash \psi$.

All that is left to show is that $\mathbb{M}''$ makes true **SND** and **CSQ**. $\mathbb{M}''$ has three further properties:

i. $\mathcal{A}''_v$ is finite for all worlds v

ii. $v R' u$ if and only if $\mathbb{M}'', u \vDash \bigwedge \mathcal{A}''(v)$

iii. The logic presented in Table 1 is sound for $\mathbb{M}''$

From (ii) and the fact that $\mathbb{M}'$ makes true (3′), we have **SND** for $\mathbb{M}''$.

All that is left is to demonstrate **CSQ** for some sound logic. We will use the logic in Table 1 itself. From (i), (ii), (iii), and the *deduction theorem* for modal logic, we have the following line of reasoning:

$$\begin{aligned}
\mathrm{Th}(\mathbb{M}'') \cup \mathcal{A}''_v \vdash \phi &\iff \mathrm{Th}(\mathbb{M}'') \vdash \bigwedge \mathcal{A}''_v \to \phi \\
&\iff \bigwedge \mathcal{A}''_v \to \phi \in \mathrm{Th}(\mathbb{M}'') \\
&\iff \mathbb{M}'', u \vDash \bigwedge \mathcal{A}''_v \to \phi \text{ for all } u \in W'' \\
&\iff \text{ for all } u \in W'', \text{ if } \mathbb{M}'', u \vDash \mathcal{A}''_v \text{ then } \mathbb{M}'', u \vDash \phi \\
&\iff \text{ for all } u \in W'', \text{ if } v R' u \text{ then } \mathbb{M}'', u \vDash \phi \\
&\iff \mathbb{M}'', v \vDash \Box \phi
\end{aligned}$$

□

Remark 5. The logic presented in Table 1 is a conservative extension of the basic modal logic K, which means that its decision problem is PSPACE hard (a lower bound). Our finitary completeness proof establishes that its complexity is in EXP2-TIME (an upper bound).

2.2 Hybrid Logic

The method of the above completeness proof makes implicit use of hybrid logic concepts. We adapt hybrid logic, first presented in [14] and then formally developed in [6], and show that it is a complete for reasoning about knowledge bases. We note that because of the increased expressive power of hybrid logic, our completeness theorem is simpler.

Definition 6. *Let Φ be a set of letters and Ψ a set of nominals, and define the language $\mathcal{L}_H(\Phi, \Psi)$ as:*

$$\phi ::= p \in \Phi \mid i \in \Psi \mid \bot \mid \phi \to \psi \mid \Box\phi \mid A{:}\,\phi \mid \forall\phi$$

Our approach in hybrid logic is to employ a universal modality along with nominals. This framework presents a logic where the agent may reason about various labeled scenarios. Scenarios may be multiply labeled or not labeled at all. From these intuitions we have the following definition:

Definition 7. *Let a **hybrid model** $\mathbb{M} = \langle W, V, R, \mathcal{A}, \ell\rangle$ be an awareness model as in Definition 2, along with a partial labeling function $\ell\colon \Psi \nrightarrow W$*

The semantics for $\vDash$ are the same as in Definition 2, only we have

$$\mathbb{M}, w \vDash \forall\phi \quad \textit{iff} \quad \forall v \in W.\mathbb{M}, v \vDash \phi$$
$$\mathbb{M}, w \vDash i \quad \textit{iff} \quad \ell(i) = w$$

The other connectives and operators are defined as usual, although we employ a few special shorthands:

Definition 8. *We employ the following special abbreviations:*

$$@_i\phi := \forall(i \to \phi) \qquad \exists\phi := \neg\forall\neg\phi \qquad \circlearrowleft_i := \Diamond i$$

This gives rise to a validity reflecting one of the axioms we saw in §2.1:

$$\vdash \circlearrowleft_i \to \Box\phi \to @_i\phi$$

This can be read as "If the agent's knowledge base is sound at world i, then if they can deduce something from it, it must be true at world i." In a way, this relativises knowledge to a world i. Worlds in epistemic logic might correspond to fantastic, pretend scenarios; however, it may be desirable for some philosophers to model knowledge about imaginary lands. After all, many children "know" a lot about Tolkien's *middle earth* or Rowling's *Hogwarts* (the school Harry Potter attends).

The semantics in Definition 7 obviates the **SND** principle we previously presented, since there is no explicit $\circlearrowleft$ symbol in this setting. **CSQ** still makes sense without modification, however. The following gives a logic for hybrid models making true **CSQ**:

$\vdash \phi \to \psi \to \phi$	$\vdash \forall(\phi \to \psi) \to \forall\phi \to \forall\psi$
$\vdash (\phi \to \psi \to \chi) \to (\phi \to \psi) \to \phi \to \chi$	$\vdash \forall\phi \to \phi$
$\vdash ((\phi \to \bot) \to (\psi \to \bot)) \to \psi \to \phi$	$\vdash \forall\phi \to \forall\forall\phi$
$\vdash \Box(\phi \to \psi) \to \Box\phi \to \Box\psi$	$\vdash \exists\phi \to \forall\exists\phi$
$\vdash A{:}\,\phi \to \Box\phi$	$\vdash \forall\phi \to \Box\phi$
	$\vdash i \to \phi \to @_i\phi$

$$\frac{\vdash \phi \to \psi \qquad \vdash \phi}{\vdash \psi} \qquad\qquad \frac{\vdash \phi}{\vdash \Box\phi} \qquad\qquad \frac{\vdash \phi}{\vdash \forall\phi}$$

Table 2. A hybrid logic for **CSQ**

Theorem 9. *Assuming an infinite store of nominals Ψ, The logic in Table 2 is sound and weakly complete for all hybrid models making true **CSQ***

Proof. As before, soundness is trivial. Completeness is based on a modification of techniques found in [5, chapter 5]. Assume that $\nvdash \psi$ and let Σ_0 be the set of subformulae of ψ, and let Υ be the set of nominals occuring in ψ. Define:

$$\Sigma_1 := \Sigma_0 \cup \{\forall\phi \mid \Box\phi \in \Sigma_0\}$$
$$\Sigma_2 := \Sigma_1 \cup \{@_i\phi\ , i \to \phi \mid \phi \in \Sigma_1\ \&\ i \in \Upsilon\}$$
$$\Sigma_3 := \Sigma_2 \cup \{\neg\phi \mid \phi \in \Sigma_2\}$$

It is easy to see that Σ_3 is finite and closed under subformulae. Now let $\mathbb{M} = \langle W, V, R_\Box, \sim_\forall, \mathcal{A}\rangle$ be the finite cannonical model formed of maximally consistent subsets of Σ_3, where everything is defined as usual, except $\sim_\forall$, which is specified as follows:

$$w \sim_\forall v \Longleftrightarrow ((\forall\phi) \in w \Longleftrightarrow (\forall\phi) \in v)$$

This model makes true the following properties:

i. $\sim_\forall$ is an equivalence relation

ii. $R_\Box \subseteq \sim_\forall$

iii. In each equivalence class specified by $\sim_\forall$, there is at most one world making true i for all $i \in \Upsilon$

The Truth Lemma for this structure is straightforward to prove. With the finitary Lindenbaum Lemma we have that some world w where $\mathbb{M}, w \nvDash \psi$. Let $\mathbb{M}' = \langle W', V', R'_\Box, \sim'_\forall, \mathcal{A}'\rangle$ be the submodel generated by $\{w\}$; we have that $\mathbb{M}', w \nvDash \psi$ [4, see chapter 2.1 for details on generated submodels]. In this model $\forall$ is a universal modality and either $[\![i]\!]^{\mathbb{M}'} = \varnothing$ or $[\![i]\!]^{\mathbb{M}'} = \{v\}$[3]. Since the store Ψ of nominals is infinite and Υ is finite, we have that $\Psi \backslash \Upsilon$ is infinite, so there is some ι: $W \hookrightarrow \Psi \backslash \Upsilon$ which assigns a fresh nominal ot every world. Let $\mathbb{M}'' = \langle W'', V'', R'', \mathcal{A}'', \ell\rangle$, where:

- $W'' := W'$
- $V''(p) := V'(p)$
- $R'' := R'_\Box$
- $\mathcal{A}''(v) := \{\phi \mid \phi \in \mathcal{A}'(v) \text{ and } \phi \text{ is a subformulae of } \psi\} \cup \{\neg\iota(u) \mid \neg v R' u\}$
- $\ell(i) := \begin{cases} w & i \in \Upsilon\ \&\ \mathbb{M}, w \vDash i \\ \iota^{-1}(i) & i \in \iota[W] \\ \text{undefined} & \text{o/w} \end{cases}$

The the worlds is $\mathbb{M}'$ and $\mathbb{M}''$ agree on all subformulae of ψ, hence $\mathbb{M}'', w \nvDash \psi$. The same reasoning as in the proof of Theorem 4 establishes and that the logic in Table 2 itself is sound for $\mathbb{M}''$ and makes true **CSQ**, establishing completeness. $\Box$

3 Logics of Multiple Knowledge Bases

So far, we have only dealt with systems that reason over a single knowledge base. In this section, we present various logics for reasoning over multiple knowledge bases.

3.1 Simplified Justification Logic

Justification Logic (JL) was originally developed by Artemov as the *Logic of Proofs* (LP) [2]. Its original purpose was to provide a framework for reasoning about explicit provability in Peano Arithmetic. The first introduction of Justification Logic can be found in [1], where Artemov and Nogina propose LP as a logic for reasoning about evidence. In [9], Melvin Fitting provided Justification Logic with Kripke model based semantics.

The innovation of LP/JL is to extend awareness logic so that awareness operations are *proof terms*. A proof term X witnesses a proposition ϕ; the notation for this is $X\colon \phi$. Proof terms are thought be operating in a *multi-conclusion* proof system, so the same proof term may witness many different propositions. Operators over proof terms are also studied. One operation of particular interest is *choice*, denoted $\oplus$. The expression "$X \oplus Y\colon \phi$" denotes that either X or Y witness ϕ. There are other operations which correspond to *modus ponens* and *proof-theoretic reflection*, however we do not consider these operations here.

In this section we reinterpret a simplified form of JL as a logic of multiple knowledge bases. This is in the spirit of JL as a logic of evidence. However, instead of reasoning about proof terms as evidence, we instead use the terms of JL for considering knowledge bases, which are taken as as corpi of evidence. We will allow for the "names" of different knowledge bases to be denoted by different terms, and terms may not refer to the same knowledge base at different worlds. This is just as how in awareness logic the agent need not be aware of the same formulae at different worlds. Finally, we will make use of JL's *choice* operator as a mechanism for forming unions of knowledge bases.

Definition 10. *Let Π be a set of primitive terms. Define $\tau(\Pi)$ with the following grammar:*

$$X ::= t \in \Pi \mid X \oplus Y$$

3. Here $[\![\phi]\!]^{\mathbb{M}}$ denotes $\{w \in W \mid \mathbb{M}, w \vDash \phi\}$

Let Φ be a set of letters and Π a set of primitive terms, and define the language $\mathcal{L}_{\mathrm{SJL}}(\Phi,\Pi)$ as:

$$\phi \ ::= \ p \in \Phi \mid \circlearrowleft_X \mid \bot \mid \phi \to \psi \mid \Box_X \phi \mid X \colon \phi$$

where $X \in \tau(\Pi)$

Definition 11. *A **simple justification model** $\mathbb{M} = \langle W, V, R, \mathcal{A}\rangle$ is a Kripke model with a valuation $V \colon \Phi \cup \{\circlearrowleft_X \mid X \in \tau(\Pi)\} \to 2^W$, an indexed relation $R \colon \tau(\Pi) \to 2^{W \times W}$, along with a modified awareness function $\mathcal{A} \colon W \times \tau(\Pi) \to 2^{\mathcal{L}_{\mathrm{SJL}}(\Phi,\Pi)}$. In practice we will denote $\mathcal{A}(w,X)$ by a curried shorthand, namely $\mathcal{A}_w(X)$.*

The semantics for $\vDash$ have the following modifications:

$$\begin{aligned} &\mathbb{M}, w \vDash \Box_X \phi \quad \textit{iff} \quad \textit{for all } v \in W \textit{ where } w R_X v \textit{ we have } \mathbb{M}, v \vDash v \\ &\mathbb{M}, w \vDash X \colon \phi \quad \textit{iff} \quad \phi \in \mathcal{A}_w(X) \end{aligned}$$

Definition 12. *The following defines properties a simply justification model may make true:*

- ***JCSQ*** — $\mathbb{M}, w \vDash \Box_X \phi$ *iff* $\mathrm{Th}(\mathbb{M}) \cup \mathcal{A}_w(X) \vdash \phi$
- ***JSND*** — $\mathbb{M}, w \vDash \circlearrowleft_X$ *iff* $\mathbb{M}, w \vDash \phi$ *for all* $\phi \in \mathcal{A}_w(X)$
- ***CHOICE*** — $\mathcal{A}_w(X \oplus Y) = \mathcal{A}_w(X) \cup \mathcal{A}_w(Y)$

*as before, $\vdash$ is any sound logical consequence relation for $\vDash$ with **modus ponens** and **reflection***

The above semantics are familiar. **JCSQ** is the same as **CSQ** from §2.1, only we now consider the theory denoted by X at w, rather than considering a single awareness function. Indeed, we may naturally see that the logic simply justification logic is a conservative extension of the awareness logic in §2.1. The awareness logic of knowledge bases is special case of simple justification logic where there is only one term.

3.2 Neighborhood Semantics

Neighborhood semantics were originally developed by Dana Scott and Richard Montague in the early 1970s as a generalization of Kripke semantics [13, 16]. In [?], Halpern and Fagin adapted neighborhood semantics for reasoning about epistemic agents without logical omniscience. In this section will demonstrate how these semantics may be modified so that every neighborhood corresponds to the logical consequences of a different knowledge base. This allows for using logics with neighborhood semantics for reasoning about multiple knowledge bases.

Definition 13. *Let Φ be a set of letters and Ψ a set of nominals, and define the language $\mathcal{L}_N(\Phi)$ as:*

$$\phi \ ::= \ p \in \Phi \mid \bot \mid \phi \to \psi \mid \Box \phi \mid A \colon \phi \mid K\phi$$

Just as previous semantics, "$\Box\phi$" is intended to mean that the agent has an argument for ϕ from some knowledge base. A novelty we present here is that "$K\phi$" means that the agent has an argument for ϕ from some *sound* knowledge base.

Definition 14. *A **neighborhood model** $\mathbb{M} = \langle W, V, \mathcal{N}, \mathcal{A}\rangle$ has a neighborhood function $\mathcal{N} \colon W \to 2^{2^W}$ and a multi-awareness function $\mathcal{A} \colon W \to 2^{2^{\mathcal{L}_N(\Phi)}}$*

The semantics for $\vDash$ have the following modifications:

$$\begin{aligned} \mathbb{M}, w \vDash \Box\phi \quad &\textit{iff} \quad \textit{there exists a } U \in \mathcal{N}_w \textit{ where } \mathbb{M}, v \vDash \phi \textit{ for all } v \in U \\ \mathbb{M}, w \vDash K\phi \quad &\textit{iff} \quad \textit{there exists a } U \in \mathcal{N}_w \textit{ where } w \in U \textit{ and } \mathbb{M}, v \vDash \phi \textit{ for all } v \in U \\ \mathbb{M}, w \vDash A \colon \phi \quad &\textit{iff} \quad \phi \in \bigcup \mathcal{A}_w \end{aligned}$$

Here we modify our previous notion of **CSQ** and **SND** to match how our semantics are intended for reasoning over multiple bases:

Definition 15. *The following defines properties a neighborhood model may make true:*

- ***NCSQ*** — $\mathbb{M}, w \vDash \Box\phi$ *iff there exists a set* $X \in \mathcal{A}_w$ *such that* $\mathrm{Th}(\mathbb{M}) \cup X \vdash \phi$
- ***NSND*** — $\mathbb{M}, w \vDash K\phi$ *iff there exists a set* $X \in \mathcal{A}_w$ *such that* $\mathbb{M}, w \vDash X$ *and* $\mathrm{Th}(\mathbb{M}) \cup X \vdash \phi$

*as before, $\vdash$ is any sound logical consequence relation for $\vDash$ with **modus ponens** and **reflection***

$$\vdash \phi \to \psi \to \phi$$
$$\vdash (\phi \to \psi \to \chi) \to (\phi \to \psi) \to \phi \to \chi$$
$$\vdash ((\phi \to \bot) \to (\psi \to \bot)) \to \psi \to \phi$$
$$\vdash K\phi \to \Box\phi$$
$$\vdash K\phi \to \phi$$
$$\vdash A{:}\, \phi \to \Box\phi$$

$$\frac{\vdash \phi \to \psi \qquad \vdash \phi}{\vdash \psi} \qquad \frac{\vdash \phi \to \psi}{\vdash \Box\phi \to \Box\psi} \qquad \frac{\vdash \phi \to \psi}{\vdash K\phi \to K\psi}$$

Table 3. A neighborhood logic for **NCSQ** and **NSND**

Theorem 16. *Assuming an infinite store of letters Φ, the logic in Table 3 is sound and weakly complete for neighborhood semantics making true **NCSQ** and **NSND**.*

Proof. As in previous proofs, we only establish completeness. Assume that $\nvdash \psi$, and define the finitary canonical model $\mathbb{M} = \langle W, V, \mathcal{K}, \mathcal{N}, \mathcal{A}\rangle$ where:

- $W :=$ the maximally consistent sets of subformulae of ψ, closed under pseudo-negation
- $V(p) := \{v \in W \mid p \in v\}$
- $\mathcal{K}_w := \{S \in \wp W \mid \{v \in W \mid \phi \in v\} \subseteq S \ \&\ K\phi \in w\}$
- $\mathcal{N}_w := \{S \in \wp W \mid \{v \in W \mid \phi \in v\} \subseteq S \ \&\ \Box\phi \in w\}$
- $\mathcal{A}_w := \{\{\phi\} \mid A{:}\, \phi \in w\}$

Note that in this model, there are neighborhoods for both knowledge and belief; more work is necessary to find a model based on this one which conforms to the semantics in Definition 14. For now, assume that K is just another modality governed by the neighborhoods in $\mathcal{K}_w$. In both neighborhood functions, our idea is that the neighborhoods of w in are supersets $S \supseteq [\![\phi]\!]$ where $\Box\phi \in w$ or $K\phi \in w$, respectively.

There are steps of the inductive proof of the Truth Lemma for this structure which are not obvious nor obviously documented elsewhere. Assume $\mathbb{M}, w \models \Box\chi$, where the Truth Lemma has been proven to hold for χ; we must show $\Box\chi \in w$. By our assumption there is some ϕ and S where $\Box\phi \in w$ and $S \supseteq \{v \in W \mid \phi \in v\}$. We argue that $\vdash \phi \to \chi$; for if not then by the finitary Lindenbaum lemma there is some $u \in W$ where $\phi \in u$ and $\chi \notin u$. Hence $u \in S$ and $\mathbb{M}, u \nvDash \chi$, a contradiction. Since $\vdash \phi \to \chi$ then we know that $\vdash \Box\phi \to \Box\chi$ by our rules, which means that $\Box\chi \in w$ by maximality.

By the Lindenbaum Lemma we have that there is some $w \in W$ where $\mathbb{M}, w \nvDash \psi$. $\mathbb{M}$ makes true a number of properties:

1. W is finite and $\mathcal{A}_v$ is finite for all $v \in W$
2. $\mathcal{K}_w \subseteq \mathcal{N}_w$
3. If $S \in \mathcal{K}_w$ then $w \in S$
4. $\mathbb{M}, v \models A{:}\, \phi \to \Box\phi$ for all $v \in W$
5. $\mathcal{N}_w$ and $\mathcal{K}_w$ are closed under supersets

As in the proof of Theorem 4, we use a bisimulation to move to a model where property (3) is strengthened to (3′): for all $S \in \mathcal{N}_w$, $w \in S$ if and only if $S \in \mathcal{K}'_w$. We defer the reader to [?, ?] for details regarding bisimulation in neighborhood semantics. Define $\mathbb{M}' := \langle W', V', \mathcal{K}', \mathcal{N}', \mathcal{A}'\rangle$ such that:

- $W' := W \uplus W$ where l, r its associated canonical injections and $\theta(v_l) := \theta(v_r) := v$ is a left-inverse of l, r
- $V'(p) := \{v_l, v_r \mid v \in V(p)\}$
- $\mathcal{K}'(v_i) := \{S \mid v_i \in S \ \&\ \theta[S] \in \mathcal{K}_v\}$ where $i = l, r$
- $\mathcal{N}'(v_i) := \mathcal{K}'(v_i) \cup \{S \mid v_i \notin S \ \&\ \theta[S] \in \mathcal{N}_v\}$ where $i = l, r$
- $\mathcal{A}'(v_l) := \mathcal{A}'(v_r) := \mathcal{A}(v)$

If we let $Z := \{(v, v_r), (v, v_l) \mid v \in W\}$, it is straightforward to verify that Z is a neighborhood bisimulation. Hence $\mathbb{M}$, $w_l \nvDash \psi$. Along with (3′), this model makes true (1), (2), (4). In fact, in light of (3′), we can see that this model does not need $\mathcal{K}'$, and conforms to the semantics in Definition 14.

As in previous proofs, let Λ be the proposition letters occuring in ψ and define an injection $\iota: W' \hookrightarrow \Phi \backslash \Lambda$, assigning a nominal to every world in $\mathbb{M}'$. Define $\mathbb{M}'' := \langle W'', V'', \mathcal{N}'', \mathcal{A}'' \rangle$ such that:

- $W'' := W'$
- $V''(p) := \begin{cases} V'(p) & p \in \Lambda \\ \{v\} & p = i(v) \\ \varnothing & o/w \end{cases}$
- $\mathcal{N}''_w := \mathcal{N}'_w$
- $\mathcal{A}''_v := \{\partial_v S \mid S \in \mathcal{N}'_w\}$

Where $\partial_v S := \{\phi \mid \phi \in \bigcup \mathcal{A}'_v, \phi \text{ is a subformula of } \psi \text{ and } S \subseteq [\![\phi]\!]^{\mathbb{M}'}\} \cup \{\neg\iota(u) \mid u \notin S\}$.
As before, this model agrees with $\mathbb{M}'$ on all subformulae of ψ, using the semantics in Definition 14, hence $\mathbb{M}'', w_l \nvDash \phi$.

All that is left is to show **NCSQ** and **NSND** three new properties:

i. For all words v and all $X \in \mathcal{A}''_v$, then X is finite and $X = \partial_v S$ for some $S \in \mathcal{N}_v$

ii. For all v and all $S \in \mathcal{N}_v$, $u \in S$ if and only if $\mathbb{M}'', u \vDash \bigwedge \partial_v S$

iii. The logic presented in Table 3 is sound for $\mathbb{M}''$

Along with the deduction theorem for $\vdash$, these properties establish **NCSQ** for this model using the logic of Table 3 itself:

$$\begin{aligned}
\exists X \in \mathcal{A}''_v \text{ s.t. } \mathrm{Th}(\mathbb{M}'') \cup X \vdash \phi &\Longleftrightarrow \exists S \in \mathcal{N}''_v \text{ s.t. } \mathrm{Th}(\mathbb{M}'') \cup \partial_v S \vdash \phi \\
&\Longleftrightarrow \exists S \in \mathcal{N}''_v \text{ s.t. } \mathrm{Th}(\mathbb{M}'') \vdash \bigwedge \partial_v S \to \phi \\
&\Longleftrightarrow \exists S \in \mathcal{N}''_v \text{ s.t. } \bigwedge \partial_v S \to \phi \in \mathrm{Th}(\mathbb{M}'') \\
&\Longleftrightarrow \exists S \in \mathcal{N}''_v \text{ s.t. } \left[\!\!\left[\bigwedge \partial_v S \right]\!\!\right] \subseteq [\![\phi]\!] \\
&\Longleftrightarrow \exists S \in \mathcal{N}''_v \text{ s.t. } \forall u \in W \text{ if } \mathbb{M}, u \vDash \bigwedge \partial_v S \text{ then } \mathbb{M}, u \vDash \phi \\
&\Longleftrightarrow \exists S \in \mathcal{N}''_v \text{ s.t. } \mathbb{M}, u \vDash \phi \text{ for all } u \in S \\
&\Longleftrightarrow \mathbb{M}, v \vDash \Box\phi
\end{aligned}$$

NSND follows from **NCSQ** and the semantics of Definition 14. □

It is not too hard to identify the fragment of this logic that governs knowledge alone. This fragment has special importance, since it governs the notion of knowledge presented in §3.3.

$$\vdash \phi \to \psi \to \phi$$
$$\vdash (\phi \to \psi \to \chi) \to (\phi \to \psi) \to \phi \to \chi$$
$$\vdash ((\phi \to \bot) \to (\psi \to \bot)) \to \psi \to \phi$$
$$\vdash K\phi \to \phi$$

$$\frac{\vdash \phi \to \psi \qquad \vdash \phi}{\vdash \psi} \qquad \frac{\vdash \phi \to \psi}{\vdash K\phi \to K\psi}$$

Table 4. A knowledge-only neighborhood logic for **NCSQ** and **NSND**

Proposition 17. *The logic in Table 4 governs the □-free fragment of the logic in Table 3*

This concludes our analysis of knowledge bases using neighborhood semantics.

3.3 Using Modalities to Quantify over Knowledge Bases

Bibliography

[1] S. Artemov and E. Nogina. Introducing justification into epistemic logic. *Journal of Logic and Computation*, 15(6):1059, 2005.

[2] S. N Artemov. Logic of proofs. *Annals of Pure and Applied Logic*, 67(1):29âĂŞ60, 1994.

[3] Kaushik Basu. *Revealed preference of government*. Cambridge University Press, Cambridge; New York, 1980.

[4] P. Blackburn, M. De Rijke, and Y. Venema. *Modal logic*. Cambridge Univ Pr, 2001.

[5] George Boolos. *The logic of provability*. Cambridge University Press, 1995.

[6] R. A. Bull. An approach to tense logic. *Theoria*, 36(3):282–300, December 1970.

[7] Donald Davidson. Could there be a science of rationality? *International Journal of Philosophical Studies*, 3(1):1âĂŞ16, 1995.

[8] R. Fagin and J. Y Halpern. Belief, awareness, and limited reasoning* 1. *Artificial Intelligence*, 34(1):39âĂŞ76, 1987.

[9] M. Fitting. The logic of proofs, semantically. *Annals of Pure and Applied Logic*, 132(1):1âĂŞ25, 2005.

[10] Joseph Y. Halpern. Set-theoretic completeness for epistemic and conditional logic. *Annals of Mathematics and Artificial Intelligence*, 26(1-4):1–27, 1999.

[11] V. F Hendricks. *Mainstream and formal epistemology*. Cambridge Univ Pr, 2006.

[12] Hilary Kornblith. Beyond foundationalism and the coherence theory. *The Journal of Philosophy*, 77(10):597–612, October 1980.

[13] Richard Montague. Universal grammar. *Theoria*, 36(3):373–398, 2008.

[14] Arthur N. Prior. *Revised and Expanded Edition of Arthur N. Prior: Papers on Time and Tense*. Oxford University Press, 1969. reprinted in 2003.

[15] Ariel Rubinstein. *Modeling Bounded Rationality*. MIT Press, 1998.

[16] D. Scott. Advice on modal logic. *Philosophical problems in Logic*, 143:74, 1970.

[17] Amartya Sen. Behaviour and the concept of preference. *Economica*, 40(159):241–259, August 1973.

[18] Johan van Benthem and F. R VelÃązquez-Quesada. Inference, promotion, and the dynamics of awareness. *ILLC Amsterdam. To appear in Knowledge, Rationality and Action*, 2009.