REPORTS
IN
INFORMATICS

ISSN 0333-3590

Only Explicitly Knowing

Thomas Ågotnes and Michal Walicki

REPORT NO 224 **February 2002**

Department of Informatics

UNIVERSITY OF BERGEN

Bergen, Norway

This report has URL http://www.ii.uib.no/publikasjoner/texrap/ps/2002-224.ps

Reports in Informatics from Department of Informatics, University of Bergen, Norway, is available at http://www.ii.uib.no/publikasjoner/texrap/.

Requests for paper copies of this report can be sent to:

Department of Informatics, University of Bergen, Høyteknologisenteret,
P.O. Box 7800, N-5020 Bergen, Norway

Abstract

It is widely accepted that there is a need for theories describing what agents explicitly know and can act upon, as opposed to what they know only implicitly. Another concept which has been suggested is called *only knowing*, and describes all that an agent knows. We argue and illustrate that a proper logic for reasoning about knowledge must combine these two concepts, and present a logic of *only explicitly knowing*. The basic semantic idea is that the agents are situated at points in the lattice of subsets of all explicit facts. While many others have set out to describe sets of explicit knowledge semantically, we introduce reasoning about such sets into the logical language. Our propositional language includes two epistemic operators: "agent i knows at least the set X" and "agent i knows at most the set X". A sound and complete logical system is presented. A additional realistic requirement is that the agents should be at *finite* points. A characterization of *finitary* sets of premises, which includes all finite sets and for which the logic is sound and complete with respect to finite agents, is presented. Incorporating additional axioms which describe purely epistemic properties corresponds to simply removing states from the set of legal epistemic states for each agent. An interesting property of our semantic model, if we require that knowledge must be true, is an incompleteness of knowledge: it is impossible to know "all I know".

Contents

Chapter 1

Introduction

It is widely accepted that there is a need for theories describing what agents *actually* or *explicitly* know and can act upon, as opposed to what they *implicitly* know, i.e. what follows logically from their explicit knowledge. This distinction is closely related to the logical omniscience problem; in classical modal epistemic logics agents know all the logical consequences of their knowledge – a description of implicit knowledge. Levesque has proposed to formalize the rules governing explicit knowledge (Levesque, 1984a), and his approach has been extended by or inspired others (see Section 7.1).

The need for reasoning about explicit knowledge is illustrated by the following example. Consider an agent a sending an encrypted version m_e of a secret message m to agent c through a public channel, that it is possible to decipher the message using two (large) prime numbers n_1 and n_2, and that the product $n = n_1 n_2$ is publicly known. Particularly, m_e and n are known by agent b. If we use the "implicit" knowledge concept, the sentence

$$\text{“agent } b \text{ knows } m\text{”} \tag{1.1}$$

could be derived from the sentences

$$\text{“agent } b \text{ knows } m_e\text{”} \tag{1.2}$$

$$\text{“agent } b \text{ knows } n\text{”} \tag{1.3}$$

assuming agent b knows the rules of arithmetic, since the values of n_1 and n_2 follows logically from the value of n.

However, even if we use the "explicit" knowledge concept, the sentence

$$\text{“agent } b \text{ does not know } m\text{”} \tag{1.4}$$

does *not* follow logically from sentences (1.2) and (1.3). Information about what the agent explicitly knows, does not make us able to deduce what he (explicitly) does *not* know. For example, agent b could have gotten to know m even before the message was sent. But if we add the sentence

$$\text{“sentences (2) and (3) describe all that is known by } b\text{”} \tag{1.5}$$

then sentence 1.4 follows from 1.2, 1.3 and 1.5. The term *only knowing* has been used by several authors (see Section 7.1) for describing all that an agent knows. However, all approaches consider only knowing in the context of *implicit* knowledge. As illustrated by our example, using the "implicit" knowledge concept does not allow us to deduce sentence 1.4 from 1.2, 1.3 and 1.5, because m is implicitly included in what is "only known" since it follows logically from other known facts. Thus, there is a need for combining reasoning about explicit knowledge with the concept of only knowing. We call this combination *only explicitly knowing*.

In this paper the implicit model of an agent is as a deliberative reasoner. The agent (explicitly) knows some formula if it has deduced it and stored it (e.g. written it on a piece of paper). We are not concerned with *how* the agents obtain their explicit knowledge, nor in the relation of this knowledge to reality (an agent can e.g. know false facts or contradictions). Particularly, we do not in general assume anything about the completeness of the deliberation mechanism, i.e. we do not assume any closure condition of explicit knowledge. In other words, our aim is, not to model the rules governing what an agent must know explicitly under given conditions, but rather to model certain interesting concepts which can be used to reason about such knowledge. We do, nonetheless, show how to model certain epistemic properties (Ch. 6).

In the next chapter, the concept of only explicitly knowing is described, together with the semantical assumptions we make. The main semantical idea is to represent an agent as a point in the subset lattice of all formulae. In Chapter 3 we define a propositional language and a semantics, capturing the assumptions, for reasoning about only explicitly knowing, while a logical system for the language is defined and shown to be sound and complete in Chapter 4. A *term calculus* is integrated in the logic, allowing us to explicitly reason about sets of formulae in the logical language. Even if we consider *explicit* knowledge, our model is not very realistic if we allow the agents to know an infinite amount of explicit facts. In Chapter 5 we add the additional epistemic (semantic) assumption that agents are finite. We show that our logical system is still complete for a certain class of theories which we call *finitary* and which includes all finite theories. In modal epistemic logic, the correspondence between the adoption of certain axioms and the properties of allowed Kripke structures has been extensively studied. In Chapter 6 we show that we can model purely epistemic axioms by simply removing certain states from the set of epistemic states an agent is allowed to be in, and discuss the epistemic properties induced by the axioms. In Section 6.4 we show that our epistemic assumptions lead to an incompleteness of knowledge: there exist true formulae which are impossible to know — more specifically, it is impossible to *know all you know*.

Chapter 2

Only Explicitly Knowing

We want to capture the concept "all an agent explicitly knows". Everything an agent *implicitly* knows may be possible to describe by one single formula α; the agent implicitly knows everything that is logically entailed by the formula. If the operator K_I denotes the concept of implicit knowledge, we can then express the agent's knowledge by $K_I\alpha$. Everything an agent *explicitly* knows, however, cannot be described by a single formula (if it knows more than one formula), because there is no closure condition for explicit knowledge. If the operator K_E denotes explicit knowledge, the formulae $K_E\alpha$ and $K_E\beta$ say that the agent knows α and β, but does not say anything about e.g. whether the agent knows the formula $\gamma = (\alpha \wedge \beta)$ (nor about whether it does *not* know the formula $\neg\alpha$). We therefore need to express knowledge about *sets* of formulae. We describe the fact that the formulae $\alpha_1, \ldots, \alpha_k$ is all that is explicitly known by agent i by the formula

$$\Diamond_i\{\alpha_1, \ldots, \alpha_k\} \tag{2.1}$$

When we consider static sets of explicit knowledge, there are a number of aspects we can reason about, for example "the agent knows more than X" or "the agent knows less than X". We formalize the fact that *at least* the formulae $\alpha_1, \ldots, \alpha_k$ is known by i by the formula

$$\triangle_i\{\alpha_1, \ldots, \alpha_k\} \tag{2.2}$$

This formula says that agent i knows each α_i, but it may know more. We use the formula

$$\bigtriangledown_i\{\alpha_1, \ldots, \alpha_k\} \tag{2.3}$$

to express the fact that agent i knows *at most* $\alpha_1, \ldots, \alpha_k$, i.e. that all he knows is included in the set but he may know less. Evidently,

$$\Diamond_i X \Leftrightarrow \triangle_i X \wedge \bigtriangledown_i X \tag{2.4}$$

2.1 Epistemic Assumptions and Knowledge Granules

The logic we construct describes groups of deliberative agents, who may possess knowledge about the world as well as about the knowledge of other agents (nested knowledge). A *deliberative* agent acts according to knowledge obtained by observing and reasoning about its environment and represented symbolically, in contrast to a *reactive* agent who reacts on stimuli from the environment without symbolical deliberation. We assume that an agent can store *pieces of knowledge*, henceforth called *knowledge granules*, which it has obtained through observation or deliberation. Initially, we make no assumption about the finiteness of this storage; this semantical assumption is added later (Ch. 5). Presently, we introduce our model of the knowledge granules. Although the set of all knowledge granules is not defined as a formal logical language, we will henceforth abuse the terminology and refer to the elements of this set as *formulae.* We assume that the representation is identical for all the agents. In general, we make no assumption about the combination of formulae in this set – i.e. we make no assumption about the consistency, or about any closure condition, of the underlying deliberation mechanism. We will, however, show how we can model such assumptions in Chapter 6. Informally, the semantics of the language we will use to reason about the agents' knowledge describes the group of agents as a collection of points in the subset lattice of all formulae. Formal semantics is presented in Section 3.2.1. We show later that certain epistemic properties of the agents can be modeled by making the set of possible points a proper subset of all points.

We define the set of knowledge granules $OL(n, \Theta)$ — just OL when no confusion can occur — over a given set Θ of primitive propositions, for a group of n agents. Θ is assumed to be countable (this assumption might be dropped which would, however, complicate some proofs).

Definition 2.1 ($\boldsymbol{OL(n,\Theta)}$): *$OL(n,\Theta)$ is the least set such that:*

- $OL_0 = \Theta$
- *If* $X \in \wp^{fin}(OL_k)$ *then* $\left.\begin{array}{c}\triangle_i X \\ \triangledown_i X\end{array}\right\} \in OL_{k+1}$

  If $\alpha, \beta \in OL_k$ *then* $\left.\begin{array}{c}\neg\alpha \\ (\alpha \wedge \beta)\end{array}\right\} \in OL_{k+1}$
- $OL(n,\Theta) = \bigcup_{k=0}^{\infty} OL_k$ $\square$

We use the common propositional connectives $\vee, \rightarrow, \leftrightarrow$ as syntactic sugar with the usual meaning, and $\Diamond_i X$ to stand for $(\triangle_i X \wedge \triangledown_i X)$.

The nesting of the power set construct, i.e., its application at each stage of the construction of OL captures the possible nesting of epistemic operators in the agents' knowledge. Examples of knowledge granules for agent i (pieces of knowledge which are held by agent i) are (if $p, q \in \Theta$ are primitive propositions): p (agent i knows p), $p \rightarrow q$ (agent i knows that p implies q), $\triangledown_j\{p\}$ (agent i knows that agent j knows at most p, and $(\triangle_j\{p \rightarrow q\} \wedge p) \rightarrow q$ (agent i knows that if agent j knows that p implies q and p is true, then q is true).

Chapter 3

Language and Semantics

A language for reasoning about explicit knowledge over OL needs a representation of the elements in OL – which again require a representation of sets of such elements. In the following, the *agent language* and the *term language* are defined by mutual recursion. The latter is simply a notation for sets of the former, while the former includes such sets in its definition.

The term language is considered first.

3.1 The Subset Lattice Term Language

The term language allows us to construct terms corresponding to the *finite* sets of formulae. The following definition is parameterized by an arbitrary agent language L, which will be instantiated in the following Section 3.2. For the moment, think of it as (something similar to) OL.

Definition 3.1 ($\boldsymbol{TL(L)}$): *Given a language L, the set of terms $TL(L)$ is the least set such that*

- *If $\alpha_1, \dots, \alpha_k \in L$ then $\underline{\{}\alpha_1\underline{,}\dots\underline{,}\alpha_k\underline{\}} \in TL(L)$*
- *If $T, U \in TL(L)$ then* $\left.\begin{array}{c}(T \sqcup U)\\(T \sqcap U)\end{array}\right\} \in TL(L)$ □

Formally, the term language $TL(L)$ is defined over the alphabet $A_L \cup \{\underline{\{}, \underline{,}, \underline{\}}, (,), \sqcup, \sqcap\}$, where A_L is the alphabet of L. We will, of course, use a term of the form $\underline{\{}\alpha_1\underline{,}\dots\underline{,}\alpha_k\underline{\}}$ to denote the set $\{\alpha_1, \dots, \alpha_k\}$. Often, to simplify the notation, we will write the term $\underline{\{}\alpha_1\underline{,}\dots\underline{,}\alpha_k\underline{\}}$ simply as $\{\alpha_1, \dots, \alpha_k\}$ (and the term $\underline{\{\}}$ as $\emptyset$). It will be clear from the context when the expression $\{\alpha_1, \dots, \alpha_k\}$ is a term, and when it is a mathematical expression of a set. Such terms are called *basic*. In the following, $\alpha, \beta, \dots$ will be used to denote formulae in the language L and $T, U, \dots$ denote general terms.

The parameter L will be omitted when it is clear from context. When defining the epistemic language in the next section, it will be instantiated once and for all.

The logical language of *term formulae* is constructed from the atoms: if T, U are terms then $T \doteq U$ is a well formed formula, and the set of all such atoms is closed under the propositional connectives $\neg$ and $\wedge$. The usual abbreviations are used, in particular, $T \preceq U$ for $T \sqcup U \doteq U$. The following definitions provide interpretation of terms and satisfaction of term formulae.

Definition 3.2: *Given a language L with the interpretation of the elements of L in OL given by a function $[\_] : L \to OL$, the interpretation of the terms $TL(L)$ is defined by the function $[\_] : TL(L) \to \wp^{fin}(OL)$ as follows:*

- $[\underline{\{}\alpha_1\underline{,}\ldots\underline{,}\alpha_n\underline{\}}] = \{[\alpha_1],\ldots,[\alpha_n]\}$
- $[S \sqcup T] = [S] \cup [T]$
- $[S \sqcap T] = [S] \cap [T]$ $\square$

Definition 3.3: *Satisfaction of term formulae is defined as expected:*

- $\models T \doteq U$ *iff* $[T] = [U]$
- $\models \neg\phi$ *iff* $\not\models \phi$*; and* $\models \phi \wedge \psi$ *iff* $\models \phi$ *and* $\models \psi$. $\square$

3.2 The Epistemic Language

We define an *agent language* $AL(n,\Theta)$ for representing formulae in $OL(n,\Theta)$, and an *epistemic language*, $EL(n,\Theta)$ which is a meta language for expressing propositions about sets of formulae in $OL(n,\Theta)$. These sets are represented by the term language $TL(AL)$. The epistemic language is a superset of the agent language and includes, in addition, term formulae over $TL(AL)$.

Definition 3.4 ($\boldsymbol{AL(n,\Theta)}$): *Given a set of primitive formulae* Θ*, the agent language* $AL(n,\Theta)$*, or just* AL*, is the least set such that:*

- $\Theta \subseteq AL$
- *If* $T \in TL(AL)$ *then* $\left.\begin{array}{c}\triangle_i T\\ \triangledown_i T\end{array}\right\} \in AL$
- *If* $\alpha,\beta \in AL$ *then* $\left.\begin{array}{c}\neg\alpha\\ (\alpha \wedge \beta)\end{array}\right\} \in AL$ $\square$

Definition 3.5 ($\boldsymbol{EL(n,\Theta)}$): *Given a set of primitive formulae* Θ*, the epistemic language* $EL(n,\Theta)$*, or just* EL*, is the least set such that:*

- $AL \subseteq EL$
- *If* $T,U \in TL(AL)$ *then* $(T \doteq U) \in EL$
- *If* $\phi,\psi \in EL$ *then* $\left.\begin{array}{c}\neg\phi\\ (\phi \wedge \psi)\end{array}\right\} \in EL$ $\square$

As usual, we use the abbreviations $(\phi \vee \psi)$ for $\neg(\neg\phi \wedge \neg\psi)$, $(\phi \rightarrow \psi)$ for $(\neg\phi \vee \psi)$, $(\phi \leftrightarrow \psi)$ for $((\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi))$ and $T \preceq U$ for $T \sqcup U \doteq U$. In addition, $\diamondsuit_i\phi$ stands for $(\triangle_i\phi \wedge \triangledown_i\phi)$. The operators $\triangle_i, \triangledown_i, \diamondsuit_i$ are called epistemic operators.

In the following, $p,q,\ldots$ are used to denote members of Θ, the meta-variables $\alpha,\beta,\ldots$ AL-formulae and $\phi,\psi,\ldots$ EL-formulae. We will not be concerned with other instantiations of L in the definition of $TL(L)$ than AL, and will often write TL for $TL(AL)$.

The set $EL(n,\Theta)$ is defined by reference to $AL(n,\Theta)$ and this latter one by a mutual recursion with the set of terms $TL(AL)$ from Definition 3.1. Intuitively, and roughly:

$TL(AL)$ consists of terms denoting *finite* sets of AL formulae with arbitrary nesting, e.g., $\{p,q,r\}$, $\{p,\triangle_i\{p,q\}\}$, $\{p,\triangle_i\{p,\triangledown_j\{q\}\}$, $\{p,\triangle_i\{q\}\} \sqcup \{\triangledown_j\{p,q,r\}\}$. A term $T \in TL(AL)$ will be interpreted as a set $[T] \in OL$.

AL consists of formulae from Θ, applications of epistemic operators to terms, and their propositional combinations, e.g., $p, \triangle_i\{p,\triangledown_j q\}, \neg\triangle_i\{p\} \wedge \triangledown_i(\{p\} \sqcup \{q,r\})$.

EL extends *AL* with propositional combinations of term formulae $T \doteq U$, $T, U \in TL(AL)$.

We include the primitive term formulae in the epistemic language in order to increase the expressiveness and make it more convenient to write down, e.g., axiom schemata. The assumption about the agent language, however, is that no reasoning about sets are necessary. In Appendix A we show that these languages defined by mutual recursion are well-defined, by showing that the corresponding recursive function has a least fixed point.

3.2.1 Semantics

The agent language is intended to represent *OL*; generally, since we can now write down a set in several ways (using different orderings of the formulae and/or using set operators), more than one formula in the agent language represents a particular knowledge granule in *OL*. We now define the interpretation of *AL*-formulae as a function $[\_] : AL(n, \Theta) \to OL(n, \Theta)$. We henceforth assume this definition in the interpretation of $TL(AL)$ (Definition 3.2).

Definition 3.6: *The interpretation of agent formulae* $[\_] : AL(n, \Theta) \to OL(n, \Theta)$ *is defined inductively:*

- $[p] = p$ *for* $p \in \Theta$
- $[\neg\alpha] = \neg[\alpha]$, *for* $\alpha \in AL(n, \Theta)$
- $[\alpha_1 \wedge \alpha_2] = [\alpha_1] \wedge [\alpha_1]$, *for* $\alpha_1, \alpha_2 \in AL(n, \Theta)$
- $[\triangle_i T] = \triangle_i[T]$, *for* $T \in TL(AL)$
- $[\triangledown_i T] = \triangledown_i[T]$, *for* $T \in TL(AL)$

where the interpretation of terms $[\_] : TL(AL) \to \wp^{fin}(OL)$ *is as in Definition 3.2.* □

We now introduce the semantical structure for which we define satisfiability of the epistemic formulae. In order to model an agent's explicit knowledge of sets of formulae, we must describe where in the subset lattice the agent "is". Thus, each agent i can be described by a state $s_i \in \wp(OL)$. In addition, we need to assign truth values to the primitive propositions. A *knowledge set structure* is a collection of points in the subset lattice $\wp(OL)$ together with a truth assignment to the primitive propositions.

Definition 3.7 (Knowledge Set Structure): *A Knowledge Set Structure (KSS) is an* $n+1$*-tuple*

$$M = (s_1, \ldots, s_n, \pi)$$

where

$$s_i \in \wp(OL) \cup \wp^{fin}(OL \cup \{*\})$$

and $\pi : \Theta \to \{\mathbf{true}, \mathbf{false}\}$ *is a truth assignment.* s_i *is called the* epistemic state *of agent* i. *The set of all knowledge set structures is denoted* $\mathcal{M}$. □

The intended interpretation of the element $*$ will be presented shortly.

At each stage of the construction of $OL(n, \Theta)$ (Def. 2.1) we construct new elements – "pieces" of knowledge – only from finite sets (of the previous stage) –

any "piece" of knowledge an agent may have is finite. $OL(\Theta)$ is the collection of such finite "pieces". An epistemic state can, however, have infinitely many (finite) pieces of knowledge. We will later be particularly interested in agents with only finite epistemic states, and will restrict the set of structures accordingly (Chapter 5). We may also want to model certain epistemic properties of the agents. Examples of this are an agent who always believes β whenever it believes both α and $(\alpha \to \beta)$, or an agent who never believes a contradiction. A result in Chapter 6 is that many interesting properties can be captured simply by restricting the set of allowed epistemic states for each agent.

As an example, the point $\{\alpha, \triangle_k\{\alpha\}, \triangle_i\{\bigtriangledown_j\{\alpha, \beta\}\}\}$ represents the epistemic state in which the agent (who "is at this point") knows: 1) α, 2) that k knows α, and 3) that i knows that j knows at most α and β.

The point s_i represents all the formulae agent i knows. Notice that an agent may "know" formulae which are not true, just as he may "know" both α and $\neg\alpha$.

Definition 3.8 (Satisfaction): *Satisfaction of a EL-formula ϕ in a KSS $M = (s_1, \ldots, s_n, \pi) \in \mathcal{M}$, written $M \models \phi$, is defined as follows:*

$$\begin{aligned}
M \models p &\Leftrightarrow \pi(p) = \textbf{true} \\
M \models \neg\phi &\Leftrightarrow M \not\models \phi \\
M \models (\phi \wedge \psi) &\Leftrightarrow M \models \phi \text{ and } M \models \psi \\
M \models \triangle_i T &\Leftrightarrow [T] \subseteq s_i \\
M \models \bigtriangledown_i T &\Leftrightarrow s_i \subseteq [T] \\
M \models T \doteq U &\Leftrightarrow [T] = [U]
\end{aligned}$$

□

Notice an important and possibly confusing subtlety. A formula which is both an AL and an EL formula, say $\phi = \triangle_i\{p, q, \bigtriangledown_j\{a, b\}\}$ is interpreted, according to Def. 3.6 as an element of $OL(n, \Theta)$ (essentially "as itself"). This, however, is merely an auxiliary definition needed for interpretation of terms. To check its satisfaction in a structure, we actually go to the point denoted by the term following the outermost epistemic operator – which is here the set $\{p, q, \bigtriangledown_j\{a, b\}\}$ – and check whether the agent i is above this point. However, if this same formula ϕ occured within a term in another formula, say $\psi = \bigtriangledown_k\{\phi\}$ then, to check the satisfaction of ψ, we would only have to check whether k is at or below the point $\{\phi\}$.

The truth condition for the derived operators are easily seen to be:

$$\begin{gathered}
M \models (\phi \vee \psi) \Leftrightarrow M \models \phi \text{ or } M \models \psi \\
M \models (\phi \to \psi) \Leftrightarrow M \not\models \phi \text{ or } M \models \psi \\
M \models \Diamond_i T \Leftrightarrow [T] = s_i \\
M \models T \preceq U \Leftrightarrow [T] \subseteq [U]
\end{gathered}$$

An *EL*-formula ϕ is *valid* with respect to a class $\mathcal{M}' \subseteq \mathcal{M}$ of knowledge set structures iff $M \models \phi$ for all $M \in \mathcal{M}'$. The model class of $\Gamma \subseteq EL$ is

$$mod(\Gamma) = \{M \in \mathcal{M} : M \models \gamma \text{ for all } \gamma \in \Gamma\}$$

A formula ϕ is a logical consequence of a set of formulae Γ, $\Gamma \models \phi$, iff ϕ is valid with respect to $mod(\Gamma)$.

Example 3.9: *Consider again the encrypted message example in the introduction (p. 3). If we assume that m_e, m, n are primitive propositions expressing the values* *of the variables in question, sentences 1.2, 1.3, 1.5 and 1.4 can be expressed as*

$$\triangle_b\{m_e\} \tag{3.1}$$

$$\triangle_b\{n\} \tag{3.2}$$

$$\triangledown_b\{m_e, n\} \tag{3.3}$$

$$\neg \triangle_b \{m\} \tag{3.4}$$

It is easy to see that the latter is a logical consequence of the three former:

$$\{\triangle_b\{m_e\}, \triangle_b\{n\}, \triangledown_b\{m_e, n\}\} \models \neg \triangle_b \{m\}$$

If the case were that we did not know whether the cunning agent b have more information than the encrypted message and the product, we would remove formula 3.3. Now, formula 3.4 does not follow from 3.1 and 3.2. □

The Element $*$

The special element $*$ is used to denote an epistemic state described by a theory (i.e. a set of formulae) Γ of a very special class of infinite theories – i.e. those in which

a) an agent cannot be below any (finite) point corresponding to a term ($\Gamma \cup \{\triangledown_i T\}$ is unsatisfiable for every T) and

b) the set of finite points the agent can be *above* is finite (there exists a "largest" term S such that $\Gamma \cup \{\triangle_i S\}$ is satisfiable and if $\Gamma \cup \{\triangle_i S'\}$ is satisfiable then $[S'] \subseteq [S]$

In other words, the agent must be at a finite point which has no corresponding term. A state s where $* \in s$ represents such a point. An example of a theory which describes such a state is the following.

Example 3.10: *Let m_e be the proposition from Example 3.9 and let*

$$\Gamma = \{\triangle_b\{m_e\}, \neg \triangledown_b \{m_e\}\} \cup \{\neg \triangle_b \{\alpha\} : \alpha \neq m_e\}$$

Γ describes a situation where

1. *Agent b knows m_e*
2. *The set $\{m_e\}$ is not all which is known by agent b*
3. *Agent b does not know any other formula than m_e*

Clearly $\Gamma \cup \{\triangle_b S\}$ is unsatisfiable for any $S \neq \{m_e\}$, so the theory Γ describes an agent with a finite state. It is also easy to see that $\Gamma \cup \{\triangledown_b T\}$ is unsatisfiable for every T – thus the agent cannot be at a finite point (denoted by a term)[1]. □

One can think of states which include the element $*$ as infinite, because they are not bounded above by any term (a) above). In Chapter 5 we discuss the case of *finite* agents, i.e. agents with states which can be described by finite terms, and will then dispense with the $*$ element.

[1]Suppose $M \models \Gamma$ where s_b is the epistemic state of agent b in M. First, let $m_e \notin [T]$. Then $s_b \nsubseteq [T]$ since $m_e \in s_b$, and thus $M \not\models \triangledown_i T$. Second, let $m_e \in [T]$. Since $M \models \neg \triangledown_b \{m_e\}$, $s_b \nsubseteq \{m_e\}$ – say $x \in s_b$ where $x \neq m_e$. If $[T] = \{m_e\}$ then $\neg \triangledown_b T \in \Gamma$ and $M \not\models \triangledown_b T$, so let $[\alpha] \in [T]$, where $[\alpha] \neq m_e$. Since $M \models \neg \triangle_b \{\alpha\}$, $[\alpha] \notin s_b$, and thus $x \neq [\alpha]$. Since $[\alpha] \neq x$ for every $[\alpha] \in [T]$, $x \notin [T]$ and because $x \in s_b$ $s_b \nsubseteq [T]$ and thus $M \not\models \triangledown_b T$.

The properties a) and b) above may seem contradictory. However, it is impossible to make the type of infinite theories such as Γ in Example 3.10 inconsistent with finite deduction rules. We use the element $*$ to provide a semantical explanation of such theories. The sound and complete logical system is presented in the next section, has only finite axioms and rules. As an alternative, we could dispense with $*$ and introduce infinite rules — a less desirable alternative[2].

[2] Let

$$R_* = \frac{\text{For all } \alpha \in AL\text{: } \Gamma \vdash \triangle_i\{\alpha\} \to \{\alpha\} \preceq T}{\Gamma \vdash \bigtriangledown_i T}$$

Let $\mathcal{M}^- = \{M \in \mathcal{M} : * \notin M\}$. In Section 4.2 we describe a logical system EC which is proven sound and complete with respect to $\mathcal{M}$. To see that $EC \cup \{R_*\}$ is sound and complete with respect to $\mathcal{M}^-$, it then suffices to show that, for any maximal consistent theory Γ, if $\Gamma \vdash \neg \bigtriangledown_i T$ for every T and $\bigcup_{\Gamma \vdash \triangle_i S}[S]$ is finite then Γ is inconsistent. Let $[S'] = \bigcup_{\Gamma \vdash \triangle_i S}[S]$, and let β be such that $\Gamma \vdash \neg\{\beta\} \preceq S'$. By soundness of the term calculus, $[\beta] \notin [S']$. By the definition of S', $\Gamma \not\vdash \triangle_i\{\beta\}$ and by maximality $\Gamma \vdash \neg \triangle_i \{\beta\}$. By R_*, $\Gamma \vdash \bigtriangledown_i S'$. Thus, since $\Gamma \vdash \neg \bigtriangledown_i S'$ by assumption, Γ is inconsistent.

Chapter 4

A Logical System

We design a logical system for our epistemic language. We use the following definitions: a logical system $S = (R, A)$ consists of a set of deduction rules R together with a set of axioms A. A formula ϕ is *provable* in S from a set of *premises* Γ, written $\Gamma \vdash_S \phi$, iff there is a finite sequence of formulae (a *proof*) where ϕ is the last one and each formula is either 1) an axiom of S, 2) a member of Γ or 3) the result of applying a deduction rule of S to a set of formulae appearing earlier in the sequence. A (*EC*) *theory* is a (possibly infinite) set of *EL*-formulae. A theory Γ is (S-) *inconsistent* iff both $\Gamma \vdash_S \phi$ and $\Gamma \vdash_S \neg\phi$ for some *EL*-formula ϕ, in which case $\Gamma \vdash_S \psi$ for *every* formula ψ. A theory Γ is *maximal* iff either $\Gamma \vdash_S \phi$ or $\Gamma \vdash \neg\phi$ for every formula ϕ. We will drop the subscript on $\vdash$ when it is clear from context.

First, in Section 4.1, we design a term calculus which is concerned exclusively with the term language. This will allow us to concentrate in the following Section 4.2 exclusively on the epistemic operators.

4.1 Term Calculus

In this section we present a *term calculus* – a logical system for propositional combinations of term equalities. It has two components. The first, the lattice calculus, is a generic axiomatization of a power set lattice. This is introduced in Section 4.1.1. The other component is an axiomatization of equality and inequality of basic terms which, again, involves interaction with the definition of the agent language. This is presented in Section 4.1.2. The term calculus is the union of these two components, and its soundness and completeness is shown in Section 4.1.3.

Recall that the language of term formulae includes term equalities $T \doteq U$ and is closed under the usual propositional connectives.

4.1.1 Lattice Calculus

We axiomatize equality in a power-set lattice.

Definition 4.1: *The (power set) lattice calculus LC is the logical system for the* *language of term formulae consisting of the following axiom schemata:*

All substitution instances of tautologies of propositional calculus		(**Prop**)
$T \doteq T$	*equivalence (reflexivity)*	(T1)
$T \doteq U \to U \doteq T$	*equivalence (symmetry)*	(T2)
$T \doteq U \wedge U \doteq V \to T \doteq V$	*equivalence (transitivity)*	(T3)
$T \doteq U \wedge S \doteq V \to S \sqcup T \doteq V \sqcup U$	*join-congruence*	(T4)
$T \doteq U \wedge S \doteq V \to S \sqcap T \doteq V \sqcap U$	*meet-congruence*	(T5)
$T \sqcup U \doteq U \sqcup T$	*join-commutativity*	(T6)
$T \sqcap U \doteq U \sqcap T$	*meet-commutativity*	(T7)
$(T \sqcup U) \sqcup V \doteq T \sqcup (U \sqcup V)$	*join-associativity*	(T8)
$(T \sqcap U) \sqcap V \doteq T \sqcap (U \sqcap V)$	*meet-associativity*	(T9)
$T \sqcup (T \sqcap U) \doteq T$	*meet-absorption*	(T10)
$T \sqcap (T \sqcup U) \doteq T$	*join-absorption*	(T11)
$T \sqcap (U \sqcup V) \doteq (T \sqcap U) \sqcup (T \sqcap V)$	*distributivity*	(T12)
$\{\alpha_1, \ldots, \alpha_n\} \doteq \{\alpha_1\} \sqcup \cdots \sqcup \{\alpha_n\}$	*atomicity*	(T13)
$\{\alpha\} \doteq \{\beta\} \to \{\alpha\} \sqcap \{\beta\} \doteq \{\alpha\}$		(T14)
$\neg(\{\alpha\} \doteq \{\beta\}) \to \{\alpha\} \sqcap \{\beta\} \doteq \emptyset$		(T15)

and the following transformation rule

$$\frac{\vdash \phi, \vdash \phi \to \psi}{\vdash \psi} \qquad \textbf{MP}$$

□

Note that in addition to the axiomatization of the distributive lattice properties (T1–T12), we also need to "connect" non-basic terms to the basic terms they (intuitively) denote. For example, we must have that $\{\alpha\} \sqcup \{\beta\} \doteq \{\alpha, \beta\}$. The axioms T13–T15 ensure this.

We now show some properties of LC. First, some well-known properties of distributive lattices (the first twelve hold for general lattices).

Lemma 4.2

1. $\vdash T \sqcup T \doteq T$
2. $\vdash T \sqcap T \doteq T$
3. $\vdash T \preceq T \sqcup U$
4. $\vdash T \sqcap U \preceq T$
5. $\vdash U \preceq T \sqcup U$
6. $\vdash T \sqcap U \preceq U$
7. $\vdash T \preceq V \wedge U \preceq V \to T \sqcup U \preceq V$
8. $\vdash V \preceq T \wedge V \preceq U \to V \preceq T \sqcap U$
9. $\vdash T \preceq V \wedge U \preceq W \to T \sqcup U \preceq V \sqcup W$
10. $\vdash T \preceq V \wedge U \preceq W \to T \sqcap U \preceq V \sqcap W$
11. $\vdash U \preceq T \leftrightarrow T \sqcap U \doteq U$
12. $\vdash T \doteq U \leftrightarrow T \preceq U \wedge U \preceq T$
13. $\vdash T \sqcup (U \sqcap V) \doteq (T \sqcup U) \sqcap (T \sqcup V)$

■

Proof See Appendix B. ■

The following Lemma justifies our previous comments regarding the set notation.

Lemma 4.3 *If the sequence $\beta_1 \ldots \beta_m$ is a permutation, possibly with duplicates, of a sequence $\alpha_1 \ldots \alpha_k$ of agent-language formulae, then*

$$\vdash \underline{\{}\alpha_1\underline{,}\ldots\underline{,}\alpha_k\underline{\}} \doteq \underline{\{}\beta_1\underline{,}\ldots\underline{,}\beta_m\underline{\}} \qquad \blacksquare$$

PROOF Follows from T13, T6 and Lemma 4.2.1. $\blacksquare$

Thus, in the notation of a basic set, the order in which we write down the formulae does not matter (with respect to equality).

4.1.2 Equality and Inequality of Basic Terms

In order to complete the term calculus, we provide an axiomatization of equality and inequality of basic terms; i.e. of agent language formulae. In short, equal formulae differ at most by the the occurrence of (syntactically) different terms which can be proven equal.

Definition 4.4 (*TC*): *The term calculus TC is the logical system for the language of term formulae, consisting of the lattice calculus LC in addition to the following axiom schemata:*

$$\begin{array}{lll}
\{\alpha_1, \ldots \alpha_k\} \preceq \{\alpha'_1, \ldots, \alpha'_m\} \to & ((\{\alpha_1\} \doteq \{\alpha'_1\} \vee \cdots \vee \{\alpha_1\} \doteq \{\alpha'_m\})\wedge & \text{N1} \\
& (\{\alpha_2\} \doteq \{\alpha'_1\} \vee \cdots \vee \{\alpha_2\} \doteq \{\alpha'_m\})\wedge & \\
& \vdots & \\
& (\{\alpha_k\} \doteq \{\alpha'_1\} \vee \cdots \vee \{\alpha_k\} \doteq \{\alpha'_m\})) & \\
\neg(\{p\} \doteq \{\alpha\}) & \textit{if } \alpha \neq p & \text{N2} \\
T \doteq S \leftrightarrow \{\triangle_i T\} \doteq \{\triangle_i S\} & & \text{N3} \\
\neg(\{\triangle_i T\} \doteq \{\alpha\}) & \textit{if } \alpha \neq \triangle_i S & \text{N4} \\
T \doteq S \leftrightarrow \{\triangledown_i T\} \doteq \{\triangledown_i S\} & & \text{N5} \\
\neg(\{\triangledown_i T\} \doteq \{\alpha\}) & \textit{if } \alpha \neq \triangledown_i S & \text{N6} \\
\{\alpha\} \doteq \{\beta\} \leftrightarrow \{\neg\alpha\} \doteq \{\neg\beta\} & & \text{N7} \\
\neg(\{\neg\alpha\} \doteq \{\beta\}) & \textit{if } \beta \neq \neg\gamma & \text{N8} \\
\{\alpha_1\} \doteq \{\beta_1\} \wedge \{\alpha_2\} \doteq \{\beta_2\} \leftrightarrow & \{\alpha_1 \wedge \alpha_2\} \doteq \{\beta_1 \wedge \beta_2\} & \text{N9} \\
\neg(\{(\alpha_1 \wedge \alpha_2)\} \doteq \{(\beta)\}) & \textit{if } \beta \neq (\beta_1 \wedge \beta_2) & \text{N10}
\end{array}$$

$\square$

N2–N10 axiomatize equality and inequality for all possible combinations of singular basic terms. N1 characterizes inequality of non-singular basic terms, in terms of inequality of singular basic terms (note that *equality* of non-singular basic terms is defined by the equivalence and congruence axioms together with the axiomatization of equivalence of singular basic terms, so the other direction of the implication N1 also holds).

4.1.3 Soundness and Completeness of Term Calculus

Lemma 4.5 (Soundness of *TC*)

$$\vdash_{TC} \phi \Rightarrow \models \phi \qquad \blacksquare$$

PROOF All axioms of the schema **Prop** are valid, since the interpretation of the propositional connectives is the same as in propositional logic. Given the interpretation of $\doteq$ as equality and $\sqcup$ and $\sqcap$ as set union and intersection, it is obvious that each of the axioms T1–T15 and N1–N10 is valid. Clearly, if $\models \phi_1$ and $\models \phi_1 \to \phi_2$ then $\models \phi_2$, so **MP** preserves validity. Let $\vdash \phi$. $\models \phi$ follows by a simple inductive proof over the length of the proof of $\vdash \phi$. $\blacksquare$

The following are some definitions and results which are used to prove completeness of the term calculus. The exposition is quite technical, and the reader may want to jump directly to the main result in Corollary 4.9 (the immediately following discussion about term terminology may, however, be useful to review).

We introduce some terminology which is useful for inductive proofs over the structure of formulae. The *degree* $d(T)$ of a term T is 1 plus the highest number of nested knowledge operators inside the term, and is defined as follows:

- $d(T \sqcup U) = \max(d(T), d(U))$
- $d(T \sqcap U) = \max(d(T), d(U))$
- $d(\{\alpha_1, \ldots, \alpha_k\}) = \max(d(\{\alpha_1\}), \ldots, d(\alpha_k\}))$
- $d(\{p\}) = 1$
- $d(\{\triangle_i S\}) = 1 + d(S)$
- $d(\{\triangledown_i S\}) = 1 + d(S)$
- $d(\{\neg\alpha\}) = d(\{\alpha\})$
- $d(\{(\alpha \wedge \beta)\}) = \max(d(\{\alpha\}), d(\{\beta\}))$

S is a subterm of a term T if it is a term and it occurs as a substring in T, and it is a proper subterm if it is different from T. A subterm S occurs at a certain *level* $l_T(S)$[1] in the term T, corresponding to the number of nested knowledge operators up to the point of the occurrence. The level of S in T is defined as follows:

$$l_T(S) = d(T) - d(S) + 1$$

As an illustration, the following term T of degree 3 is annotated by the level in T (above) and degree (below) of each subterm:

$$\overbrace{\underbrace{\{\triangle_i \overbrace{\underbrace{\{p, \triangledown_i \overbrace{\underbrace{\overbrace{\underbrace{\{q\}}_{1}}^{3} \sqcap \overbrace{\underbrace{\{q,r\}}_{1}}^{3}}_{1}}^{3}\}}_{2}}^{2}, \triangledown_i \overbrace{\underbrace{\{\triangle_j \overbrace{\underbrace{\{p\}}_{1}}^{3}\}}_{2}}^{2}\}}_{3}}^{1}$$

In the following lemmata, we use the notation $S[\cdot]$ and $\phi[\cdot]$ for a formula and a term, respectively, with a *hole*, i.e. with a *missing term.* If T is a term, then $S[T]$ ($\phi[T]$) is a well-formed term (formula).

A hole can occur at any place in a term or formula where a well-formed term can occur, hence a hole occurs at a certain level in the term or formula – and it may occur at *any* level.

We now show that we can replace a subterm with an *equal* subterm, at an arbitrary level inside a term.

[1] Of course, the same subterm may occur at different levels in the term. Formally, we can view each occurrence of a subterm as being annotated by a unique name. In other words, "the level of S in T is the level of a particular occurrence of S in T.

Lemma 4.6 (Substitution) *The following transformation rule is admissible in TC:*

$$\frac{\vdash T \doteq U}{\vdash S[T] \doteq S[U]} \qquad \textbf{Repl}$$

■

Proof We first show an intermediate result; that we can substitute equal terms on the first level, i.e. that the following rule is admissible:

$$\text{If } S[\cdot] \text{ has a hole at level 1 then } \frac{\vdash T \doteq U}{\vdash S[T] \doteq S[U]} \qquad \textbf{Repl}^-$$

Terms with holes at level 1 can be defined in the following way. $[\cdot]$ is a term with a hole at level 1, and if $S[\cdot]$ is a term with a hole at level 1 and V is a term then $S[\cdot] \sqcup V$, $V \sqcup S[\cdot]$, $S[\cdot] \sqcap V$, $V \sqcap S[\cdot]$ are terms with holes at level 1. The proof is by structural induction over $S[\cdot]$. Assume $\vdash T \doteq U$. In the base case, $S[\cdot]$ is empty and $S[T] = T$, and the consequent of the transformation rule is the same as the antecedent. For the induction step, $S[\cdot]$ is constructed by combining a formula with a hole with a well-formed formula using $\sqcup$ or $\sqcap$. We first consider the case where $S[\cdot] = S_1[\cdot] \sqcup S_2$, i.e. that the hole in S is in the left sub term:

1	$\vdash T \doteq U$	Assumption
2	$\vdash S_1[T] \doteq S_1[U]$	Ind. hyp.
3	$\vdash S_2 \doteq S_2 \wedge S_1[T] \doteq S_1[U] \rightarrow S_1[T] \sqcup S_2 \doteq S_1[U] \sqcup S_2$	T4
4	$\vdash S_2 \doteq S_2$	T1
5	$\vdash S_2 \doteq S_2 \wedge S_1[T] \doteq S_1[U]$	**MP**(**Prop**,2,4)
6	$\vdash S_1[T] \sqcup S_2 \doteq S_1[U] \sqcup S_2$	**MP**(5,3)

We next consider $S[\cdot] = S_1[\cdot] \sqcap S_2$:

1	$\vdash T \doteq U$	Assumption
2	$\vdash S_1[T] \doteq S_1[U]$	Ind. hyp.
3	$\vdash S_2 \doteq S_2 \wedge S_1[T] \doteq S_1[U] \rightarrow S_1[T] \sqcap S_2 \doteq S_1[U] \sqcap S_2$	T5
4	$\vdash S_2 \doteq S_2$	T1
5	$\vdash S_2 \doteq S_2 \wedge S_1[T] \doteq S_1[U]$	**MP**(**Prop**,2,4)
6	$\vdash S_1[T] \sqcap S_2 \doteq S_1[U] \sqcap S_2$	**MP**(5,3)

The proofs in the cases where $S[\cdot] = S_1 \sqcup S_2[\cdot]$ and $S[\cdot] = S_1 \sqcap S_2[\cdot]$ are similar. This completes the proof of **Repl**$^-$.

The proof of **Repl** is by induction over the level of the hole in $S[\cdot]$. Let $\vdash T \doteq U$. For the induction base, consider a term $S[\cdot]$ with a hole in level 1. Then $\vdash S[T] \doteq S[U]$ by **Repl**$^-$.

For the inductive step, assume that **Repl** holds for all terms with holes at level k. Let $S[\cdot]$ be a term with a hole at level $k+1$. Then $S[\cdot]$ contains a level k term on the form $V[\cdot] = \{\alpha_1, \ldots, \alpha_j[\cdot], \ldots, \alpha_m\}$, where $\alpha_j[\cdot]$ contains a term with a hole on level 1. We now show that

$$\vdash \{\alpha_j[T]\} \doteq \{\alpha_j[U]\} \tag{4.1}$$

by induction over the structure of $\alpha_j[\cdot]$. For the base case, first let $\alpha_j[\cdot] = \triangle_i W[\cdot]$ where $W[\cdot]$ is a term with a missing level 1 subterm. By **Repl**$^-$, $\vdash W[T] \doteq W[U]$, and thus $\vdash \{\triangle_i W[T]\} \doteq \{\triangle_i W[U]\}$ by N3. Second, let $\alpha_j[\cdot] = \triangledown_i W[\cdot]$, and $\vdash \{\triangledown_i W[T]\} \doteq \{\triangledown_i W[U]\}$ follows similarly by **Repl**$^-$ and N5. For the induction step, first let $\alpha_j[\cdot] = \neg\beta[\cdot]$. By the induction hypothesis, $\vdash \{\beta[T]\} \doteq \{\beta[U]\}$ and thus $\vdash \{\neg\beta[T]\} \doteq \{\neg\beta[U]\}$ by N7. Second, let $\alpha_j[\cdot] = (\beta_1[\cdot] \wedge \beta_2)$. By the induction hypothesis $\vdash \{\beta_1[T]\} \doteq \{\beta_1[U]\}$, and by T1 $\vdash \{\beta_2\} \doteq \{\beta_2\}$, and thus $\vdash \{(\beta_1[T] \wedge \beta_2)\} \doteq \{(\beta_1[U] \wedge \beta_2)\}$ by N9. Third, let $\alpha_j[\cdot] = (\beta_1 \wedge \beta_2[\cdot])$ and $\vdash \{(\beta_1 \wedge \beta_2[T])\} \doteq \{(\beta_1 \wedge \beta_2[U])\}$ by similar reasoning. This completes the proof of (4.1). Then, it follows from several applications of T13 and T4 that $\vdash V[T] \doteq V[U]$. If we replace the term $V[\cdot]$ in $S[\cdot]$ with a hole, we get a term $S'[\cdot]$ with a missing k level term. From $\vdash V[T] \doteq V[U]$ and the induction hypothesis, we have that $\vdash S'[V[T]] \doteq S'[V[U]]$. Clearly, $S'[V[T]] = S[T]$ and $S'[V[U]] = S[U]$ and thus we have that $\vdash S[T] \doteq S[U]$ which completes the inductive step in the main proof. ■

Completeness of the term calculus follows as a corollary of the next lemma, the proof of which is quite involved and found in an appendix.

Lemma 4.7 *Let T_1 and T_2 be terms.*

1. *If $[T_1] = [T_2]$ then $\vdash T_1 \doteq T_2$*
2. *If $[T_1] \neq [T_2]$ then $\vdash \neg T_1 \doteq T_2$* ■

PROOF See Appendix C. ■

Corollary 4.8 *Let T and U be terms.*

1. *If $[T] \subseteq [U]$ then $\vdash T \preceq U$*
2. *If $[T] \not\subseteq [U]$ then $\vdash \neg T \preceq U$* ■

PROOF If $[T] \subseteq [U]$ then $[T \sqcup U] = [T] \cup [U] = [U]$, and $\vdash T \sqcup U \doteq U$ by Lemma 4.7.1. If $[T] \not\subseteq [U]$ then $[T \sqcup U] = [T] \cup [U] \neq [U]$, and $\vdash \neg T \sqcup U \doteq U$ by Lemma 4.7.2. ■

Finally, we show completeness of the term calculus.

Corollary 4.9 *If ϕ is a term formula, then*

$$\models \phi \Longrightarrow \vdash \phi$$ ■

PROOF We prove the following statement by structural induction over ϕ:

$$\text{If } \models \phi \text{ then } \vdash \phi, \text{ otherwise } \vdash \neg\phi \tag{4.2}$$

In the base case, $\phi = T \doteq U$. If $\models \phi$ then $[T] = [U]$ and $\vdash T \doteq U$ by Lemma 4.7.1. If $\not\models \phi$ then $[T] \neq [U]$ and $\vdash \neg T \doteq U$ by Lemma 4.7.2. For the induction step, first assume that $\phi = \neg\psi$. If $\models \neg\psi$ then $\not\models \psi$ and $\vdash \neg\psi$ by the induction hypothesis. If $\not\models \neg\psi$ then $\models \psi$, $\vdash \psi$ by the induction hypothesis and $\vdash \neg\neg\psi$ by **Prop**. Second, assume that $\phi = (\psi_1 \wedge \psi_2)$. If $\models \phi$, $\models \psi_1$ and $\models \psi_2$, $\vdash \psi_1$ and $\vdash \psi_2$ by the induction hypothesis and $\vdash (\psi_1 \wedge \psi_1)$ by **Prop**. If $\not\models \phi$ then $\not\models \psi_1$ or $\not\models \psi_2$, $\vdash \neg\psi_1$ or $\vdash \neg\psi_2$ by the induction hypothesis, and thus $\vdash \neg(\psi_1 \wedge \psi_2)$ by **Prop**. ■

4.2 Epistemic Calculus

Finally, we introduce a logical system for the epistemic language. The epistemic language includes the language of term formulae, and the epistemic calculus includes the term calculus.

Definition 4.10 (*EC*): *The epistemic calculus EC is the logical system for the epistemic language EL consisting of the term calculus TC in addition to the following axiom schemata:*

All substitution instances of tautologies of propositional calculus	**Prop**
$\triangle_i \emptyset$	E1
$(\triangle_i T \wedge \triangle_i U) \rightarrow \triangle_i (T \sqcup U)$	E2
$(\triangledown_i T \wedge \triangledown_i U) \rightarrow \triangledown_i (T \sqcap U)$	E3
$(\triangle_i T \wedge \triangledown_i U) \rightarrow T \preceq U$	E4
$(\triangledown_i (U \sqcup \{\alpha\}) \wedge \neg \triangle_i \{\alpha\}) \rightarrow \triangledown_i U$	E5
$\triangle_i T \wedge U \preceq T \rightarrow \triangle_i U$	**KS**
$\triangledown_i T \wedge T \preceq U \rightarrow \triangledown_i U$	**KG**

and the following transformation rule

$$\frac{\Gamma \vdash \phi, \Gamma \vdash \phi \rightarrow \psi}{\Gamma \vdash \psi} \qquad \textbf{MP}$$

□

Example 4.11: *Consider again the encrypted message example in the introduction (p. 3). In Example 3.9 (p. 10) we used the following formulae to represent sentences 1.2, 1.3, 1.5 and 1.4:*

$$\triangle_b \{m_e\}$$
$$\triangle_b \{n\}$$
$$\triangledown_b \{m_e, n\}$$
$$\neg \triangle_b \{m\}$$

We show that:

$$\{\triangle_b \{m_e\}, \triangle_b \{n\}, \triangledown_b \{m_e, n\}\} \vdash \neg \triangle_b \{m\}$$

Let $\Gamma = \{\triangle_b \{m_e\}, \triangle_b \{n\}, \triangledown_b \{m_e, n\}\}$.

1	$\Gamma \vdash ((\triangle_b \{m\} \wedge \triangledown_b \{m_e, n\}) \rightarrow \{m\} \preceq \{m_e, n\})$ $\rightarrow ((\triangledown_b \{m_e, n\} \wedge \neg \{m\} \preceq \{m_e, n\}) \rightarrow \neg \triangle_b \{m\})$	***Prop***
2	$\Gamma \vdash (\triangle_b \{m\} \wedge \triangledown_b \{m_e, n\}) \rightarrow \{m\} \preceq \{m_e, n\}$	*E4*
3	$\Gamma \vdash (\triangledown_b \{m_e, n\} \wedge \neg \{m\} \preceq \{m_e, n\}) \rightarrow \neg \triangle_b \{m\}$	***MP****(2,3)*
4	$\Gamma \vdash \neg(\{m\} \preceq \{m_e, n\})$	*Lemma 4.8.2*
5	$\Gamma \vdash \triangledown_b \{m_e, n\}$ $\rightarrow (\neg \{m\} \preceq \{m_e, n\} \rightarrow (\triangledown_b \{m_e, n\} \wedge \neg \{m\} \preceq \{m_e, n\}))$	***Prop***
6	$\Gamma \vdash \triangledown_b \{m_e, n\}$	*Premise*
7	$\Gamma \vdash \neg \{m\} \preceq \{m_e, n\} \rightarrow (\triangledown_b \{m_e, n\} \wedge \neg \{m\} \preceq \{m_e, n\})$	***MP****(5,6)*
8	$\Gamma \vdash \triangledown_b \{m_e, n\} \wedge \neg \{m\} \preceq \{m_e, n\}$	***MP****(4,7)*
9	$\Gamma \vdash \neg \triangle_b \{m\}$	***MP****(3,8)* □

In Example 4.11 we explicitly wrote down every deduction in the proof, except in line (4). We will henceforth often drop trivial details in the proofs, by e.g. combining several applications of **Prop** and **MP** in one line. Since the set of term formulae (from the previous section) is a subset of *EL*, we will, as in line (4) above, refer to the completeness of the term calculus instead of writing out the proofs of term (in)equalities.

The deduction theorem, $\Gamma \cup \{\gamma\} \vdash \phi \Rightarrow \Gamma \vdash \gamma \rightarrow \phi$, clearly holds for *EC*, and will be referred to as *DT*.

4.2.1 Soundness and Completeness

We show that EC is sound and complete[2].

Lemma 4.12 (Soundness)

$$\textit{If } \Gamma \vdash \phi \textit{ then } \Gamma \models \phi \qquad \blacksquare$$

PROOF We show that all axioms of EC are valid (we only need to consider the "new" axioms in Definition 4.10; all the axioms of TC are valid since TC is sound). Let $M = (s_1, \ldots, s_n, \pi)$ be an arbitrary knowledge set structure. We show that $M \models \phi$ for each axiom ϕ.

Prop All **Prop**-axioms are valid, since the interpretation of the propositional connectives is the same as in propositional logic.

E1 $\phi = \triangle_i \emptyset$. $[\emptyset] = \emptyset \subseteq s_i \Rightarrow M \models \phi$

E2 $\phi = (\triangle_i T \wedge \triangle_i U) \rightarrow \triangle_i (T \sqcup U)$. If $M \not\models \triangle_i T \wedge \triangle_i U$, then $M \models \phi$ trivially. If $M \models \triangle_i T \wedge \triangle_i U$, then

$$\left.\begin{array}{l} M \models \triangle_i T \Rightarrow [T] \subseteq s_i \\ M \models \triangle_i U \Rightarrow [U] \subseteq s_i \end{array}\right\} \Rightarrow [T \sqcup U] = [T] \cup [U] \subseteq s_i \Rightarrow M \models \triangle_i (T \sqcup U)$$

and thus $M \models \phi$.

E3 $\phi = (\triangledown_i T \wedge \triangledown_i U) \rightarrow \triangledown_i (T \sqcap U)$. If $M \not\models \triangledown_i T \wedge \triangledown_i U$, then $M \models \phi$ trivially. If $M \models \triangledown_i T \wedge \triangledown_i U$, then

$$\left.\begin{array}{l} M \models \triangledown_i T \Rightarrow s_i \subseteq [T] \\ M \models \triangledown_i U \Rightarrow s_i \subseteq [U] \end{array}\right\} \Rightarrow s_i \subseteq [T] \cap [U] = [T \sqcap U] \Rightarrow M \models \triangledown_i (T \sqcap U)$$

and thus $M \models \phi$.

E4 $\phi = (\triangle_i T \wedge \triangledown_i U) \rightarrow T \preceq U$. If $M \not\models \triangle_i T \wedge \triangledown_i U$, then $M \models \phi$ trivially. If $M \models \triangle_i T \wedge \triangledown_i U$, then

$$\left.\begin{array}{l} M \models \triangle_i T \Rightarrow [T] \subseteq s_i \\ M \models \triangledown_i U \Rightarrow s_i \subseteq [U] \end{array}\right\} \Rightarrow [T] \subseteq [U] \Rightarrow M \models T \preceq U$$

and thus $M \models \phi$.

KS $\phi = (\triangle_i T \wedge U \preceq T) \rightarrow \triangle_i U$. If $M \not\models \triangle_i T \wedge U \preceq T$, then $M \models \phi$ trivially. If $M \models \triangle_i T \wedge U \preceq T$, then

$$\left.\begin{array}{l} M \models \triangle_i T \Rightarrow [T] \subseteq s_i \\ M \models U \preceq T \Rightarrow [U] \subseteq [T] \end{array}\right\} \Rightarrow [U] \subseteq s_i \Rightarrow M \models \triangle_i U$$

and thus $M \models \phi$.

KG $\phi = (\triangledown_i T \wedge T \preceq U) \rightarrow \triangledown_i U$. If $M \not\models \triangledown_i T \wedge T \preceq U$, then $M \models \phi$ trivially. If $M \models \triangledown_i T \wedge T \preceq U$, then

$$\left.\begin{array}{l} M \models \triangledown_i T \Rightarrow s_i \subseteq [T] \\ M \models T \preceq U \Rightarrow [T] \subseteq [U] \end{array}\right\} \Rightarrow s_i \subseteq [U] \Rightarrow M \models \triangledown_i U$$

and thus $M \models \phi$.

[2] Soundness and completeness of EC relies on soundness and completeness of TC. As a generalization, *any* sound and complete calculus of term (in)equality could be used instead of TC.

Clearly, if $\Gamma \models \phi_1$ and $\Gamma \models \phi_1 \to \phi_2$ then $\Gamma \models \phi_2$, so **MP** preserves logical consequence.

Let $\Gamma \vdash \phi$. $\Gamma \models \phi$ follows by a simple inductive proof over the length of the proof of $\Gamma \vdash \phi$. ∎

In order to show completeness, we first show the counterpart of Lemma 4.6 for *EC* — we can substitute equals for equals inside an *EL*-formula.

Lemma 4.13 (Substitution) *The following transformation rule is admissible in EC:*

$$\frac{\Gamma \vdash T \doteq U}{\Gamma \vdash \phi[T] \leftrightarrow \phi[U]} \qquad \textbf{Subst}$$

∎

Proof (**Subst**) The proof is by structural induction over $\phi[\cdot]$[3]. The base cases are the three types of atomic formulae involving a term. In the first base case, $\phi[\cdot] = \triangle_i S[\cdot]$:

1	$\Gamma \vdash T \doteq U$	Assumption
2	$\Gamma \vdash S[T] \doteq S[U]$	**Repl**(1)
3	$\Gamma \vdash S[U] \sqcup S[T] \doteq S[T] \sqcup S[T]$	**Repl**(2)
4	$\Gamma \vdash S[T] \sqcup S[T] \doteq S[T]$	Lemma 4.2.1
5	$\Gamma \vdash S[U] \sqcup S[T] \doteq S[T]$	T3(3,4)
6	$\Gamma \vdash S[U] \preceq S[T]$	Def. $\preceq$ (5)
7	$\Gamma \vdash \triangle_i S[T] \wedge S[U] \preceq S[T] \to \triangle_i S[U]$	**KS**
8	$\Gamma \vdash (\triangle_i S[T] \wedge S[U] \preceq S[T] \to \triangle_i S[U])$	
	$\to (S[U] \preceq S[T] \to (\triangle_i S[T] \to \triangle_i S[U]))$	**Prop**
9	$\Gamma \vdash \triangle_i S[T] \to \triangle_i S[U]$	**MP**(7,6,5)

The proof for $\Gamma \vdash \triangle_i S[U] \to \triangle_i S[T]$ can be obtained by using T2 with line 1, and exchanging $S[T]$ and $S[U]$ in the rest of the proof above. In the second base case, $\phi[\cdot] = \triangledown_i S[\cdot]$, and the proof is obtained by replacing the three last lines of the preceding proof with:

7	$\Gamma \vdash \triangledown_i S[T] \wedge S[U] \preceq S[T] \to \triangledown_i S[U]$	**KG**
8	$\Gamma \vdash (\triangledown_i S[T] \wedge S[U] \preceq S[T] \to \triangledown_i S[U])$	
	$\to (S[U] \preceq S[T] \to (\triangledown_i S[T] \to \triangledown_i S[U]))$	**Prop**
9	$\Gamma \vdash \triangledown_i S[T] \to \triangledown_i S[U]$	**MP**(7,6,5)

and similarly for $\Gamma \vdash \triangledown_i S[U] \to \triangledown_i S[T]$.

We now consider the third base case where $\phi[\cdot] = (S[\cdot] \doteq V)$ (the proof in the case where $\phi[\cdot] = (S \doteq V[\cdot])$ is similar):

1	$\Gamma \vdash T \doteq U$	Assumption
2	$\Gamma \vdash S[U] \doteq S[T]$	**Repl**(1)
3	$\Gamma \vdash S[U] \doteq S[T] \wedge S[T] \doteq V \to S[U] \doteq V$	T3
4	$\Gamma \vdash S[U] \doteq S[T] \to (S[T] \doteq V \to S[U] \doteq V)$	**MP**,**Prop**, 3
5	$\Gamma \vdash S[T] \doteq V \to S[U] \doteq V$	**MP**,2,4

The proof of $\Gamma \vdash S[U] \doteq V \to S[T] \doteq V$ can be obtained by using T2 with line (1), and exchanging $S[T]$ and $S[U]$ in the rest of the proof above.

[3]Formulae and terms with holes were discussed in Section 4.1.3.

In the inductive step, first consider the case where $\phi[\cdot] = \neg\psi[\cdot]$. By the induction hypothesis, $\Gamma \vdash \psi[T] \leftrightarrow \psi[U]$ whenever $\Gamma \vdash T \doteq U$, and $\Gamma \vdash \neg\psi[T] \leftrightarrow \neg\psi[U]$ follows by **Prop**. Second, consider the case where $\phi[\cdot] = \psi_1[\cdot] \wedge \psi_2$ (the proof in the case that $\phi[\cdot] = \psi_1 \wedge \psi_2[\cdot]$ is similar). By the induction hypothesis, $\Gamma \vdash \psi_1[T] \leftrightarrow \psi_1[U]$ whenever $\Gamma \vdash T \doteq U$, and $\Gamma \vdash \psi_1[T] \wedge \psi_2 \leftrightarrow \psi_1[U] \wedge \psi_2$ by **Prop**. Clearly, $\psi_1[T] \wedge \psi_2 = (\psi_1 \wedge \psi_2)[T]$ (the hole in $(\psi_1 \wedge \psi_2)[\cdot]$ is the hole in ψ_1), so $\Gamma \vdash (\psi_1 \wedge \psi_2)[T] \leftrightarrow (\psi_1 \wedge \psi_2)[U]$. ■

We now show completeness of *EL*, with respect to $\mathcal{M}$, by the common strategy of showing satisfiability of maximal consistent theories. It may be instructive to recall that for every finite set $X \in \wp^{fin}(OL)$, there exists a term T such that $[T] = X$.

Lemma 4.14 *If Γ is a consistent theory, then there exists a maximal consistent theory $\Gamma' \supseteq \Gamma$.* ■

Proof We construct $\Gamma' \supseteq \Gamma$ as follows. *EL* is countable, so let $\phi_1, \phi_2, \ldots$ be an enumeration of positive formulae from *EL* (i.e., those where the outermost operator is not $\neg$):

$$\Gamma_0 = \Gamma \tag{4.3}$$

$$\Gamma_{i+1} = \begin{cases} \Gamma_i \cup \{\phi_{i+1}\} & \text{if } \Gamma_i \cup \{\phi_{i+1}\} \text{ is consistent} \\ \Gamma_i \cup \{\neg\phi_{i+1}\} & \text{otherwise} \end{cases} \tag{4.4}$$

$$\Gamma' = \bigcup_{i=0}^{\infty} \Gamma_i \tag{4.5}$$

Since Γ_0 is consistent, clearly all Γ_i are consistent. Let Δ be a finite subset of Γ'. $\Delta \subseteq \Gamma_j$ for some j, so Δ must be consistent. Since all finite subsets of Γ' are consistent, so is Γ'. Since either $\phi \in \Gamma'$ or $\neg\phi \in \Gamma'$ for any ϕ, Γ' is maximal. ■

Lemma 4.15 *Let Γ' be a maximal consistent theory. If there exists a T' such that $\Gamma' \vdash \triangledown_i T'$, then there exists a T such that $\Gamma' \vdash \Diamond_i T$.* ■

Proof Let $\Gamma' \vdash \triangledown_i T'$. Let T be a term such that

$$[T] = \bigcap_{[S] \subseteq [T'] \text{ and } \Gamma' \vdash \triangledown_i S} [S]$$

(this set is finite since it is included in $[T']$, and thus such a term T exists). Since $[T']$ is finite it has only a finite number of subsets, and thus $\Gamma' \vdash \triangledown_i T$ can be obtained by a finite number of applications of E3, completeness of the term calculus and **Subst**. Let U be a term such that

$$[U] = \bigcup_{\Gamma' \vdash \triangle_i S} [S]$$

To see that this set indeed is finite, and thus that the term U exists, observe that for any term S such that $\Gamma' \vdash \triangle_i S$, $[S] \subseteq [T]$ by E4 and soundness of the term calculus, so there can only be a finite number of such terms. By a finite number of applications of E2, completeness of the term calculus and **Subst**, $\Gamma' \vdash \triangle_i U$. By E4 we then have that $\Gamma' \vdash U \preceq T$ and thus $[U] \subseteq [T]$ by soundness of the term calculus.

We now show that $[T] \subseteq [U]$. Assume the opposite; that there is a x such that $x \in [T]$ and $x \notin [U]$. Then there is a α such that $[\alpha] = x$. Let T^- be a term such that

$$[T^-] = [T] \setminus [\{\alpha\}]$$

(T^- must exist since $[T]$ is finite). $[T] = [T^- \sqcup \{\alpha\}]$ and by completeness of the term calculus $\Gamma' \vdash T \doteq T^- \sqcup \{\alpha\}$. Since $\Gamma' \vdash \bigtriangledown_i T$, $\Gamma' \vdash \bigtriangledown_i (T^- \sqcup \{\alpha\})$ by **Subst**. If $\Gamma' \vdash \triangle_i\{\alpha\}$ then $[\alpha] \in U$ which is a contradiction, hence $\Gamma' \vdash \neg \triangle_i \{\alpha\}$ by the maximality of Γ'. By E5, $\Gamma' \vdash \bigtriangledown_i T^-$. But then $[T] \subseteq [T^-]$, which is a contradiction. Therefore, $[T] \subseteq [U]$. Thus $[T] = [U]$ and, by completeness of the term calculus, $\Gamma' \vdash T \doteq U$ and, by **Subst**, $\Gamma' \vdash \Diamond_i T$. ∎

Lemma 4.16 *Every maximal consistent theory is satisfiable.* ∎

Proof Let Γ be a maximal consistent theory. We construct a knowledge set structure

$$M^\Gamma = (s_1^\Gamma, \ldots, s_2^\Gamma, \pi^\Gamma) \tag{4.6}$$

where

$$\pi^\Gamma(p) = \mathbf{true} \Leftrightarrow \Gamma \vdash p \tag{4.7}$$

$$s_i^\Gamma = \begin{cases} [W] \text{ where } \Gamma \vdash \Diamond_i W & \text{if there is a } T' \text{ such that } \Gamma \vdash \bigtriangledown_i T' \\ \bigcup_{\Gamma \vdash \triangle_i S}[S] \cup \{*\} & \text{otherwise} \end{cases} \tag{4.8}$$

intended to satisfy Γ. In the definition of s_i^Γ, any W such that $\Gamma \vdash \Diamond_i W$ can be chosen in the case that there exists a T' such that $\Gamma \vdash \bigtriangledown_i T'$; Lemma 4.15 guarantees the existence of such a term. We show, by structural induction over ϕ, that

$$M^\Gamma \models \phi \Longleftrightarrow \Gamma \vdash \phi \tag{4.9}$$

This is a stronger statement than the lemma; the lemma is given by the direction to the left.

For the base case:

- $\phi = p \in \Theta$: $\Gamma \vdash p$ iff $\pi^\Gamma(p) = \mathbf{true}$ iff $M^\Gamma \models p$.
- $\phi = T \doteq U$: (4.9) follows by soundness and completeness of TC.
- $\phi = \triangle_i T$: For each direction of (4.9), we consider the cases where

  a) there is a T' such that $\Gamma \vdash \bigtriangledown_i T'$ or

  b) $\Gamma \nvdash \bigtriangledown_i T'$ for every T'

  (the two cases in the definition of s_i^Γ).

  $\Rightarrow$) If $M^\Gamma \models \triangle_i T$ then $[T] \subseteq s_i^\Gamma$.

  a) $[T] \subseteq [W]$, and $\Gamma \vdash T \preceq W$ by completeness of TC. Since $\Gamma \vdash \triangle_i W$, $\Gamma \vdash \triangle_i T$ by **KS**.

  b) Let $[T] = \{[\alpha_1], \ldots, [\alpha_k]\}$ ($[T] \in \wp^{fin}(OL)$). By the definition of s_i^Γ there exists, for $1 \leq j \leq k$, a proof $\Gamma \vdash \triangle_i T_{\alpha_j}$ where T_{α_j} is a term such that $[\alpha_j] \in [T_{\alpha_j}]$ (since $* \notin [S]$ for any term S). Since $[\{\alpha_j\}] \subseteq [T]$, $\Gamma \vdash \{\alpha_j\} \preceq T_{\alpha_j}$ $(1 \leq j \leq k)$ by completeness of the term calculus and then $\Gamma \vdash \triangle_i\{\alpha_j\}$ by **KS**. By (repeated applications of) E2, $\Gamma \vdash \triangle_i(\{\alpha_1\} \sqcup \cdots \sqcup \{\alpha_k\})$. Since $[\{\alpha_1\} \sqcup \cdots \sqcup \{\alpha_k\}] = [T]$, $\Gamma \vdash \triangle_i T$ by completeness of the term calculus and **Subst**.

  $\Leftarrow$) **a)** Since $\Gamma \vdash \bigtriangledown_i W$, if $\Gamma \vdash \triangle_i T$ then $\Gamma \vdash T \preceq W$ by E4. If $[T] \not\subseteq [W]$ then $\Gamma \vdash \neg T \preceq W$ by completeness of TC and Γ would be inconsistent, so $[T] \subseteq [W]$ and thus $M^\Gamma \models \triangle_i T$.

  b) If $\Gamma \vdash \triangle_i T$ then $[T] \subseteq s_i^\Gamma$ and thus $M^\Gamma \models \triangle_i T$.

- $\phi = \bigtriangledown_i T$: For each direction of (4.9), we again consider the cases where

  a) there is a T' such that $\Gamma \vdash \bigtriangledown_i T'$ and

  b) $\Gamma \nvdash \bigtriangledown_i T'$ for every T'

  (the two cases in the definition of s_i^Γ).

  $\Rightarrow$) If $M^\Gamma \models \bigtriangledown_i T$ then $s_i^\Gamma \subseteq [T]$.

  a) $[W] \subseteq [T]$, and $\Gamma \vdash W \preceq T$ by completeness of TC. Because $\Gamma \vdash \bigtriangledown_i W$, $\Gamma \vdash \bigtriangledown_i T$ by **KG**.

  b) If $s_i^\Gamma \subseteq [T]$ then $* \in [T]$, which is not the case for any T. This case is impossible.

  $\Leftarrow$) **a)** Since $\Gamma \vdash \bigtriangleup_i W$, if $\Gamma \vdash \bigtriangledown_i T$ then $\Gamma \vdash W \preceq T$ by E4. If $[W] \nsubseteq [T]$ then $\Gamma \vdash \neg W \preceq T$ by completeness of TC and Γ would be inconsistent, so $[W] \subseteq [T]$ and thus $M^\Gamma \models \bigtriangledown_i T$.

  b) Since $\Gamma \nvdash \bigtriangledown_i T$ for any T, this case is impossible.

For the induction step:

- $\phi = \neg\psi$: $M^\Gamma \models \neg\psi$ iff $M^\Gamma \not\models \psi$ iff (by the induction hypothesis) $\Gamma \nvdash \psi$ iff (by the fact that Γ is both consistent and maximal) $\Gamma \vdash \neg\psi$.

- $\phi = (\psi_1 \wedge \psi_2)$: $M^\Gamma \models (\psi_1 \wedge \psi_2)$ iff $M^\Gamma \models \psi_1$ and $M^\Gamma \models \psi_2$ iff (by the induction hypothesis) $\Gamma \vdash \psi_1$ and $\Gamma \vdash \psi_2$ iff (by **Prop**) $\Gamma \vdash (\psi_1 \wedge \psi_2)$. ∎

Lemma 4.17 *Every consistent theory is satisfiable.* ∎

PROOF Follows immediately from Lemmas 4.14 and 4.16. ∎

Theorem 4.18 *If $\Gamma \models \phi$ then $\Gamma \vdash \phi$.* ∎

PROOF Let $\Gamma \models \phi$. Assume that Γ is consistent (otherwise $\Gamma \vdash \phi$ trivially). Then, by Lemma 4.17, there exists a M such that $M \models \Gamma$ and thus $M \models \phi$. $M \not\models \neg\phi$, and $\Gamma \cup \{\neg\phi\}$ is unsatisfiable and thus inconsistent. Then, $\Gamma \cup \{\neg\phi\} \vdash \phi$, $\Gamma \vdash \neg\phi \rightarrow \phi$ by DT, and $\Gamma \vdash \phi$ by **Prop**. ∎

Chapter 5

Finite Agents and Finitary Theories

The motivation behind describing the concept of explicit knowledge is the idea that it is different from implicit knowledge; it is unrealistic to assume that an agent would deduce all possible consequences — generally infinitely many — of his knowledge. In fact, it is not only unrealistic to assume that an agent could *deduce* an infinite amount of facts, but also that he is able to explicitly know (i.e. to store) an infinite amount of facts at all. The latter would require an infinite amount of memory. In the current section, we add the semantical assumption that agents know only a finite amount of facts at a given time. Such agents are henceforth called finite agents.

Semantically, we model finite agents by only considering the class of structures $\mathcal{M}_{fin}$ where each epistemic state is finite[1]:

$$\mathcal{M}_{fin} = \{(s_1, \ldots, s_n, \pi) \in \mathcal{M} : s_i \in \wp^{fin}(OL)\}$$

The only way to describe an agent with an *infinite* epistemic state in *EL*, is to use an infinite set of formulae. Thus, for finite sets of premises *EC* is complete with respect to $\mathcal{M}_{fin}$. But for certain infinite sets of premises which "place" an agent at an infinite point, *EC* is incomplete with respect to $\mathcal{M}_{fin}$ and, furthermore, there is no way to extend *EC* properly without using transformation rules with infinitely many premises. We now define a set of theories, called *finitary*, for which *EC* is complete with respect to $\mathcal{M}_{fin}$ (*EC* is of course sound with respect to $\mathcal{M}_{fin} \subseteq \mathcal{M}$ (also) for finitary theories).

5.1 Finitary Theories

Definition 5.1 (Finitary Theory): *A theory* Γ *is* finitary *iff it is consistent and for all* ϕ,[2]

$$(\Gamma \vdash (\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n) \rightarrow \phi \textit{ for all terms } T_1, \ldots, T_n) \Rightarrow \Gamma \vdash \phi \qquad \square$$

Before we describe the intuition behind finitary theories, we present one more definition and a few observations.

[1] In addition to the infinite states we dispense with the finite states which include the $*$ element. Recall from the discussion in Section 3.2.1 that we view these states as infinite, since they cannot be described by any terms.

[2] In higher order notation: Γ is finitary iff it is consistent and $\forall\phi((\forall T_1 \cdots \forall T_n \Gamma \vdash (\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n) \rightarrow \phi) \Rightarrow \Gamma \vdash \phi)$.

Definition 5.2 (Finitarily Open Theory): *A theory* Γ *is* finitarily open *iff there exist terms* $T_1, \ldots, T_n$ *such that*

$$\Gamma \not\vdash \neg(\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n) \qquad \square$$

Lemma 5.3 *1. A finitary theory is finitarily open.*

2. If Γ *is a finitary theory and* $\Gamma \not\vdash \phi$*, then* $\Gamma \cup \{\neg\phi\}$ *is finitarily open.* ■

PROOF 1. Let Γ be a finitary theory. If Γ is not finitarily open, $\Gamma \vdash \neg(\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n)$ for all terms $T_1, \ldots, T_n$. Then, for an arbitrary ϕ, $\Gamma \vdash (\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_i T_n) \to \phi$ for all $T_1, \ldots, T_n$ and thus $\Gamma \vdash \phi$ since Γ is finitary. By the same argument $\Gamma \vdash \neg\phi$, contradicting the fact that Γ is consistent.

2. Let Γ be a finitary theory, and let $\Gamma \not\vdash \phi$. Then there must exist terms $T_1^\phi, \ldots, T_n^\phi$ such that $\Gamma \not\vdash (\triangledown_1 T_1^\phi \wedge \cdots \wedge \triangledown_n T_n^\phi) \to \phi$. By **Prop** we must have that $\Gamma \not\vdash \neg\phi \to \neg(\triangledown_1 T_1^\phi \wedge \cdots \wedge \triangledown_n T_n^\phi)$ and thus that $\Gamma \cup \{\neg\phi\} \not\vdash \neg(\triangledown_1 T_1^\phi \wedge \cdots \wedge \triangledown_n T_n^\phi)$, which shows that $\Gamma \cup \{\neg\phi\}$ is finitarily open. ■

Corollary 5.4 *Let Cons, FOT and Fin be the set of all consistent, finitarily open and finitary theories, respectively.*

$$Fin \subset FOT \subset Cons \qquad ■$$

PROOF Follows immediately (see Example 5.5 below for examples showing that these subsets are proper.). ■

The following is an argument for why we must require theories for finite agents to be finitary. First, the concept of a *finitarily open* theory requires that it is consistent to "place" each of the n agents at a finite point simultaneously. In other words, a theory which is not finitarily open describes *infinite* agents – agents which cannot be at finite points. However, to secure that $\Gamma \vdash \phi$ whenever $\Gamma \models \phi$ (completeness), it is not enough to require that Γ is finitarily open – we must also require that $\Gamma \cup \{\neg\phi\}$ is finitarily open whenever $\Gamma \not\vdash \phi$. Otherwise $\Gamma \cup \{\neg\phi\}$ is unsatisfiable in $\mathcal{M}_{fin}$ (it describes one or more agents at infinite points) and then $\Gamma \models \phi$ which contradicts completeness.

Example 5.5: *The following are examples of non-finitary theories (let* $n = 2$ *and* $p \in \Theta$*):*

1. $\Gamma_1 = \{\triangle_2\{p\}, \triangle_2\{\triangle_2\{p\}\}, \triangle_2\{\triangle_2\{\triangle_2\{p\}\}\}, \ldots\}$. Γ_1 *is not finitarily open, and describes an agent with an infinite epistemic state.*

2. $\Gamma_2 = \{\neg \triangledown_1 T : T \in TL\}$. Γ_2 *is not finitarily open, and describes an agent which cannot be at any finite point.*

3. $\Gamma_3 = \{\triangledown_1 T \to \neg \triangledown_2 T' : T, T' \in TL\}$. Γ_3 *is not finitarily open, and describes a situation where agents 1 and 2 cannot* simultaneously *be at finite points.*

4. $\Gamma_4 = \{\triangledown_1 T \to p : T \in TL\}$. Γ_4 *is finitarily open, but not finitary. To see the former, observe that there is no* T *such that* $\Gamma_4 \vdash \neg \triangledown_1 T$ *or* $\Gamma_4 \vdash \neg \triangledown_2 T$*. To see the latter, observe that* $\Gamma_4 \not\vdash p$ *but* $\Gamma_4 \vdash (\triangledown_1 T_1 \wedge \triangledown_2 T2) \to p$ *for all* T_1, T_2*.*

$\square$

5.2 Completeness

We show that EC is complete with respect to $\mathcal{M}_{fin}$, for all finitary theories.

Definition 5.6 (Maximal Finitarily Open Theory (MFOT)): *A theory* Γ *is* a maximal finitarily open theory *(MFOT) iff it is maximal and finitarily open.* $\square$

Lemma 5.7 *If* Γ *is a finitarily open theory, then there exists a MFOT* $\Gamma' \supseteq \Gamma$. $\blacksquare$

Proof Let Γ be a finitarily open theory, and let $T_1, \ldots, T_n$ be such that $\Gamma \not\vdash \neg(\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n)$. We construct Γ' as follows. EL is countable, so let $\phi_1, \phi_2, \ldots$ be an enumeration of positive formulae from EL (i.e., those where the outermost operator is not $\neg$):

$$\Gamma_0 = \Gamma \cup \{\triangledown_1 T_1, \ldots, \triangledown_n T_n\} \tag{5.1}$$

$$\Gamma_{i+1} = \begin{cases} \Gamma_i \cup \{\phi_{i+1}\} & \text{if } \Gamma_i \cup \{\phi_{i+1}\} \text{ is consistent} \\ \Gamma_i \cup \{\neg\phi_{i+1}\} & \text{otherwise} \end{cases} \tag{5.2}$$

$$\Gamma' = \bigcup_{i=0}^{\infty} \Gamma_i \tag{5.3}$$

We first show that Γ' is consistent. Γ is consistent. Assume that $\Gamma \cup \{\triangledown_1 T_1, \ldots, \triangledown_n T_n\}$ is inconsistent. Then $\Gamma \cup \{\triangledown_1 T_1, \ldots, \triangledown_n T_n\} \vdash \neg(\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n)$ and thus $\Gamma \vdash \neg(\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n)$ by DT and **Prop**, which is not the case. Thus, Γ_0 is consistent. Since Γ_0 is consistent, clearly all Γ_i are consistent. Let Δ be a finite subset of Γ'. $\Delta \subseteq \Gamma_j$ for some j, so Δ must be consistent. Since all finite subsets of Γ' are consistent, so is Γ'.

Clearly $\Gamma' \vdash (\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n)$, and by consistency $\Gamma' \not\vdash \neg(\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n)$ and Γ' is finitarily open.

Since either $\phi \in \Gamma'$ or $\neg\phi \in \Gamma'$ for any ϕ, Γ' is maximal. $\blacksquare$

Lemma 5.8 *Every MFOT* Γ *is satisfiable in* $\mathcal{M}_{fin}$. $\blacksquare$

Proof Let Γ be a MFOT, and let $T_1, \ldots, T_n$ be such that $\Gamma \not\vdash \neg(\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n)$. Since Γ is maximal $\Gamma \vdash (\triangledown_1 T_1 \wedge \cdots \wedge \triangledown_n T_n)$, by **Prop** $\Gamma \vdash \triangledown_i T_i$ $(1 \le i \le n)$, and by Lemma 4.15 (Γ is finitarily open and thus consistent) there exist T_i' $(1 \le i \le n)$ such that $\Gamma \vdash \Diamond_i T_i'$. We construct a knowledge set structure

$$M_f^\Gamma = (s_1^\Gamma, \ldots, s_n^\Gamma \pi^\Gamma) \tag{5.4}$$

where

$$\pi^\Gamma(p) = \mathbf{true} \Leftrightarrow \Gamma \vdash p \tag{5.5}$$

$$s_i^\Gamma = [T_i'] \tag{5.6}$$

intended to satisfy Γ. We show, by structural induction over ϕ that

$$M_f^\Gamma \models \phi \Longleftrightarrow \Gamma \vdash \phi \tag{5.7}$$

This is a stronger statement than the lemma; the lemma is given by the direction to the left.

For the base case:

- $\phi = p \in \Theta$: $\Gamma \vdash p$ iff $\pi^\Gamma(p) = \mathbf{true}$ iff $M_f^\Gamma \models p$.
- $\phi = T \doteq U$: (5.7) follows by soundness and completeness of TC.

- $\phi = \triangle_i T$:
  - $\Rightarrow$) If $M_f^\Gamma \models \triangle_i T$ then $[T] \subseteq [T_i']$ and $\Gamma \vdash T \preceq T_i'$ by completeness of TC. Since $\Gamma \vdash \triangle_i T_i'$, $\Gamma \vdash \triangle_i T$ by **KS**.
  - $\Leftarrow$) Since $\Gamma \vdash \triangledown_i T_i'$, if $\Gamma \vdash \triangle_i T$ then $\Gamma \vdash T \preceq T_i'$ by E4. By soundness of TC, $[T] \subseteq [T_i']$, so $M_f^\Gamma \models \triangle_i T$.
- $\phi = \triangledown_i T$:
  - $\Rightarrow$) If $M_f^\Gamma \models \triangledown_i T$ then $[T_i'] \subseteq [T]$ and $\Gamma \vdash T_i' \preceq T$ by completeness of TC. Since $\Gamma \vdash \triangledown_i T_i'$, $\Gamma \vdash \triangledown_i T$ by **KG**.
  - $\Leftarrow$) Since $\Gamma \vdash \triangle_i T_i'$, if $\Gamma \vdash \triangledown_i T$ then $\Gamma \vdash T_i' \preceq T$ by E4. By soundness of TC, $[T_i'] \subseteq [T]$, so $M_f^\Gamma \models \triangledown_i T$.

For the induction step:

- $\phi = \neg\psi$: $M_f^\Gamma \models \neg\psi$ iff $M_f^\Gamma \not\models \psi$ iff (by the induction hypothesis) $\Gamma \nvdash \psi$ iff (by the fact that Γ is both consistent and maximal) $\Gamma \vdash \neg\psi$.
- $\phi = (\psi_1 \wedge \psi_2)$: $M_f^\Gamma \models (\psi_1 \wedge \psi_2)$ iff $M_f^\Gamma \models \psi_1$ and $M_f^\Gamma \models \psi_2$ iff (by the induction hypothesis) $\Gamma \vdash \psi_1$ and $\Gamma \vdash \psi_2$ iff (by **Prop**) $\Gamma \vdash (\psi_1 \wedge \psi_2)$. ■

Corollary 5.9 *Every finitarily open theory Γ is satisfiable in $\mathcal{M}_{fin}$.* ■

Proof Follows immediately from Lemmas 5.7 and 5.8. ■

Theorem 5.10 *For every finitary theory Γ, if $\Gamma \models \phi$ with respect to $\mathcal{M}_{fin}$ then $\Gamma \vdash \phi$.* ■

Proof Let Γ be a finitary theory. In the case that Γ is inconsistent, $\Gamma \vdash \phi$ trivially so let Γ be consistent. Let $\Gamma \models \phi$. By Lemma 5.3.1 Γ is finitarily open and thus satisfiable by Corollary 5.9. $\Gamma \cup \{\neg\phi\}$ is unsatisfiable, and thus not finitarily open. Assume that $\Gamma \nvdash \phi$. Then $\Gamma \cup \{\neg\phi\}$ is finitarily open by Lemma 5.3.2. The contradiction shows that the assumption must be wrong, and that $\Gamma \vdash \phi$. ■

Soundness and completeness of modal epistemic logic is traditionally presented as the fact that the set of theorems is equal to the set of valid formulae, i.e. that $\vdash \phi \Leftrightarrow \models \phi$. Syntactical consequence can then be defined as $\Gamma \vdash \phi$ iff there is a finite set $\{\psi_1, \ldots, \psi_k\} \subseteq \Gamma$ such that $\vdash (\psi_1 \rightarrow (\psi_2 \rightarrow \cdots \phi) \cdots)$. It is clear then that $\Gamma \vdash \phi \Leftrightarrow \Gamma \models \phi$. This gives completeness only for *finite* theories, since we can only use a finite number of premises. The fact that completeness is restricted to finitary theories in our system, should be seen in the light of the situation in traditional epistemic logic.

5.3 Axiomatization Using Infinite Rules

The reason that the logic EC is complete only for finitary theories, is that we cannot axiomatize the fact that certain infinite sets of formulae lead to inconsistency using only finite deduction rules. Here, we describe our logic using *infinite* deduction rules (and thus infinite proofs) and show that it is sound and complete for all theories.

The infinite rule we need is

$$R_f = \frac{\text{for all } T \in TL\text{: } \Gamma \cup \{\triangledown_i T\} \vdash \bot}{\Gamma \vdash \bot}$$

(where $\bot$ is an arbitrary contradiction $\bot = \phi \wedge \neg\phi$).

We now argue briefly that $EC \cup \{R_f\}$ is sound and complete with respect to $\mathcal{M}_{fin}$, for all theories (most of the proof is identical to parts of the two previous soundness and completeness proofs, so we refer to those).

5.3.1 Soundness

For soundness we only have to show that R_f preserves logical consequence. Let $\Gamma \cup \{\triangledown_i T\} \models \perp$ for every term T, and assume that $M \models \Gamma$ where s_i is the epistemic state of i in M. Then $M \not\models \{\triangledown_i T\}$, and thus $s_i \not\subseteq [T]$ for every term T. Since there is a term T for every set $s \in \wp^{fin}(OL)$, this is a contradiction. Thus, Γ has no models and $\Gamma \models \perp$.

5.3.2 Completeness

We can extend every consistent theory Γ to a maximal consistent theory Γ' by using the same construction as in Lemma 4.14. Since Γ is consistent, there is a term T' such that $\Gamma \cup \{\triangledown_i T'\}$ is consistent (since otherwise R_f could be used to show inconsistency of Γ). By construction of Γ', $\triangledown_i T' \in \Gamma'$. Assume that $\Gamma' \vdash \perp$. If this proof is finite, R_f was not applied and we have a contradiction by compactness (see Lemma 4.14). If the proof of $\Gamma' \vdash \perp$ is infinite, then R_f must have been applied. $\Gamma' \cup \{\triangledown_i T'\} \vdash \perp$ must have been among the premises in this application, but this is equal to the conclusion $\Gamma' \vdash \perp$. It is thus clear that we cannot prove the conclusion using R_f, and we cannot have an infinite proof either. Thus, Γ' is consistent (and clearly maximal).

Clearly, Lemma 4.15 still holds for $EC \cup \{R_f\}$. The fact that every maximal consistent theory are satisfiable in $\mathcal{M}_{fin}$, can now be shown in exactly the same way as in Lemma 5.8; since the construction of the maximal theory (above) guarantees that $\triangledown_i T' \in \Gamma'$ we can "place" the agent at $[T']$: $s_i^\Gamma = [T']$.

It follows that every consistent theory is satisfiable, and thus that $EC \cup \{R_f\}$ is complete (for all theories) by the same argument as in Theorem 4.18.

Chapter 6

Epistemic Properties

Epistemic properties of the individual agents can be modeled by imposing constraints on the belief state sets. Ideally, such semantic constraints should correspond to the acceptance of certain axioms. In modal epistemic logic, the correspondence between axioms and semantical constraints has been extensively studied. In this chapter, we first define the notion of an *epistemic axiom* — an axiom which can be used to describe an purely epistemic property of an agent. We then show that the model class of an epistemic axiom can be constructed by removing illegal epistemic states. Finally, we investigate the epistemic properties induced by some axioms commonly used in modal epistemic logic. Like modal epistemic logics extended with certain axioms are sound and complete with respect to a constrained class of Kripke structures, our logic extended with these axioms are sound and complete with respect to the constructed model class.

6.1 Axioms and Model Class Construction

Given a set of axioms Φ, we will construct a class of structures $\mathcal{M}^\Phi$ such that

$$\mathcal{M}^\Phi = mod(\Phi)$$

First, we construct the model class $\mathcal{M}^\phi$ for a single formula ϕ:

Definition 6.1 ($\mathcal{M}^\phi$): *For each formula $\phi \in EL$, $\mathcal{M}^\phi \subseteq \mathcal{M}$ is defined by structural induction over ϕ as follows:*

$$\begin{aligned}
\phi = p \in \Theta : &\quad \mathcal{M}^\phi = \{(s_1, \ldots, s_n, \pi) \in \mathcal{M} : \pi(p) = \mathbf{true}\} \\
\phi = \triangle_i T : &\quad \mathcal{M}^\phi = \{(s_1, \ldots, s_n, \pi) \in \mathcal{M} : [T] \subseteq s_i\} \\
\phi = \bigtriangledown_i T : &\quad \mathcal{M}^\phi = \{(s_1, \ldots, s_n, \pi) \in \mathcal{M} : s_i \subseteq [T]\} \\
\phi = \neg\psi : &\quad \mathcal{M}^\phi = \mathcal{M} \backslash \mathcal{M}^\psi \\
\phi = \psi_1 \wedge \psi_2 : &\quad \mathcal{M}^\phi = \mathcal{M}^{\psi_1} \cap \mathcal{M}^{\psi_2}
\end{aligned}$$

□

It is easy to see that the corresponding construction of model classes of formulae including the derived operators is done by taking:

$$\begin{aligned}
\phi = \Diamond_i T : &\quad \mathcal{M}\phi = \{(s_1, \ldots, s_n, \pi) \in \mathcal{M} : s_i = [T]\} \\
\phi = \psi_1 \vee \psi_2 : &\quad \mathcal{M}^\phi = \mathcal{M}^{\psi_1} \cup \mathcal{M}^{\psi_2} \\
\phi = \psi_1 \rightarrow \psi_2 = \neg\psi_1 \vee \psi_2 : &\quad \mathcal{M}^\phi = \mathcal{M} \backslash \mathcal{M}^{\psi_1} \cup \mathcal{M}^{\psi_2}
\end{aligned}$$

The construction in Def. 6.1 is a straightforward application of the rules of the semantics, and it is trivial to show that $\mathcal{M}^\phi$ in fact is the model class of ϕ:

Lemma 6.2

$$\mathcal{M}^{\phi} = mod(\phi) \qquad \blacksquare$$

Proof We show that

$$M \models \phi \Leftrightarrow M \in \mathcal{M}^{\phi}$$

where $M = (s_1, \ldots, s_n, \pi)$ by structural induction over ϕ:

$\phi = p \in \Theta$: $M \models \phi \Leftrightarrow \pi(p) = \textbf{true} \Leftrightarrow M \in \mathcal{M}^{\phi}$

$\phi = \triangle_i T$: $M \models \phi \Leftrightarrow [T] \subseteq s_i \Leftrightarrow M \in \mathcal{M}^{\phi}$

$\phi = \triangledown_i T$: $M \models \phi \Leftrightarrow s_i \subseteq [T] \Leftrightarrow M \in \mathcal{M}^{\phi}$

$\phi = \neg\psi$: By the induction hypothesis, $M \models \psi \Leftrightarrow M \in \mathcal{M}^{\psi}$. $M \models \phi \Leftrightarrow M \not\models \psi \Leftrightarrow M \notin \mathcal{M}^{\psi} \Leftrightarrow M \in \mathcal{M}^{\phi}$.

$\phi = \psi_1 \wedge \psi_2$: By the induction hypothesis, $M \models \psi_1 \Leftrightarrow M \in \mathcal{M}^{\psi_1}$ and $M \models \psi_2 \Leftrightarrow M \in \mathcal{M}^{\psi_2}$. $M \models \phi \Leftrightarrow M \models \psi_1$ and $M \models \psi_2 \Leftrightarrow M \in \mathcal{M}^{\psi_1}$ and $M \in \mathcal{M}^{\psi_2} \Leftrightarrow M \in \mathcal{M}^{\phi}$. $\blacksquare$

When constructing the model classes, the usual care must be taken to specify whether we want to describe *an axiom* (a formula) or an *axiom schema* (a set of formulae). As a sentence which can be interpreted as both, consider

$$\triangle_i T \rightarrow \triangle_i U$$

When this sentence is specified as an axiom, we should construct the model class $\mathcal{M}^{\phi}$ taking the unspecified sets T, U and the agent name i as parameters:

$$S_i^{\phi} = \wp(OL) \backslash \{X \in \wp(OL) : [T] \subseteq X\} \cup \{X \in \wp(OL) : [U] \subseteq X\}$$

When specified as an axiom schema, we should try to construct the model class for all formulae Φ satisfying the schema. In the current case,

$$\mathcal{M}^{\Phi} = \emptyset$$

(Of course, we may want to fix only *some* of the parameters in an axiom schema).

The model class of a set of formulae Φ is, of course, the intersection of the model classes for the individual formulae:

$$\mathcal{M}^{\Phi} = \bigcap_{\phi \in \Phi} \mathcal{M}^{\phi} \tag{6.1}$$

When we extend the semantics by restricting the class of allowed knowledge set structures to $\mathcal{M}' \subset \mathcal{M}$, logical consequence with respect to $\mathcal{M}'$ is denoted $\Gamma \models_{\mathcal{M}'} \phi$ and is defined usual: for every $M \in \mathcal{M}'$, if $M \models \Gamma$ then $M \models \phi$.

6.2 Epistemic Axioms

The most probable reason for extending the system *EC* by adding new axioms — or, equivalently, impose semantical constraints by restricting the set of allowed structures — is that we want to model certain epistemic properties of the agents, e.g. the fact that they will never "know" a contradiction. Not all formulae in *EL* should be considered as candidates for describing epistemic properties. One example is $p \to \triangle_i\{p\}$. This formula does not solely describe the *agent* – it describes a relationship between the agent and the world, i.e. about an agent in a situation. Another example is $\Diamond_i\{p\} \to \Diamond_j\{q\}$, which describes a constraint on one agent's belief set contingent on another agent's belief set. Neither of these two formulae describes *epistemic* properties of an agent. Candidates for *epistemic axioms* should therefore a) only refer to epistemic facts and not to external facts and b) only describe one particular agent. We will call formulae satisfying a) *epistemic* formulae. In the following definition, EF is the set of epistemic formulae and Ax is the set of candidate epistemic axioms.

Definition 6.3 (EF, EF^i, Ax):

- *$EF \subseteq EL$ is the least set such that*

$$\begin{array}{rll} \text{If } T \in TL \text{ then} & \triangle_i T, \triangledown_i T \in EF & (1 \leq i \leq n) \\ \text{If } \phi, \psi \in EF \text{ then} & \neg\phi, (\phi \wedge \psi) \in EF & \end{array}$$

- *$EF^i = \{\phi \in EF$: Every epistemic operator in ϕ is an i-operator$\}$ $(1 \leq i \leq n)$*
- *$Ax = \bigcup_{1 \leq i \leq n} EF^i$* □

6.2.1 Model Class Construction for Epistemic Axioms

In Section 6.1, we defined the construction of the model class of a formula. However, this definition does not define any useful properties of the model class. Contrast this with modal epistemic logic, which for certain interesting axioms has model classes with properties like "all reflexive structures" and "all transitive structures" (more on this in Section 6.3). We now show that for the class of epistemic axioms we can restrict the set of legal epistemic states *locally*; i.e. for each axiom $\phi \in Ax$ and agent i we can construct a set S_i^ϕ of legal epistemic states. In other words, incorporating epistemic axioms corresponds to removing illegal epistemic states of single agents.

We now show that, in the case that $\phi \in Ax$, we can rephrase Def. 6.4 in terms of sets of legal epistemic states for each agent.

Definition 6.4 ($\mathcal{M}^\phi_{Ax}$): *For each epistemic formula $\phi \in EF^i$,*

$$\mathcal{M}^\phi_{Ax} = \{(s_1^\phi, \ldots, s_n^\phi, \pi^\phi) \in \mathcal{M} : s_i \in S_i^\phi\}$$

where S_i^ϕ is constructed by structural induction over ϕ as follows:

$$\begin{array}{rl} \phi = \triangle_i T : & S_i^\phi = \{X \in \wp(OL) : [T] \subseteq X\} \\ \phi = \triangledown_i T : & S_i^\phi = \{X \in \wp(OL) : X \subseteq [T]\} \\ \phi = \neg\psi : & S_i^\phi = \wp(OL) \backslash S_i^\psi \\ \phi = \psi_1 \wedge \psi_2 : & S_i^\phi = S_i^{\psi_1} \cap S_i^{\psi_2} \end{array}$$ □

In the construction of $\mathcal{M}^\phi_{Ax}$ we remove the impossible epistemic states by restricting the set of epistemic states to S_i^ϕ. The epistemic states which are not removed are the possible states — an agent can be placed in any of these states and will satisfy the epistemic axiom ϕ.

Again, it is easy to see that the corresponding construction of model classes of formulae including the derived operators is done by taking:

$$\begin{array}{rl} \phi = \Diamond_i T : & S_i^\phi = \{[T]\} \\ \phi = \psi_1 \vee \psi_2 : & S_i^\phi = S_i^{\psi_1} \cup S_i^{\psi_2} \\ \phi = \psi_1 \to \psi_2 = \neg\psi_1 \vee \psi_2 : & S_i^\phi = \wp(OL) \backslash S_i^{\psi_1} \cup S_i^{\psi_2} \end{array}$$

Lemma 6.5

$$\mathcal{M}^{\phi}_{Ax} = mod(\phi) \qquad \blacksquare$$

PROOF We show that

$$M \models \phi \Leftrightarrow M \in \mathcal{M}^{\phi}_{Ax}$$

where $M = (s_1, \ldots, s_n, \pi)$ by structural induction over $\phi \in EF^i$:

$\phi = \triangle_i T$: $M \models \phi \Leftrightarrow [T] \subseteq s_i \Leftrightarrow s_i \in S_i^{\phi} \Leftrightarrow M \in \mathcal{M}^{\phi}_{Ax}$

$\phi = \triangledown_i T$: $M \models \phi \Leftrightarrow s_i \subseteq [T] \Leftrightarrow s_i \in S_i^{\phi} \Leftrightarrow M \in \mathcal{M}^{\phi}_{Ax}$

$\phi = \neg\psi$: By the induction hypothesis, $M \models \psi \Leftrightarrow M \in \mathcal{M}^{\psi}_{Ax}$. $M \models \phi \Leftrightarrow M \not\models \psi \Leftrightarrow M \notin \mathcal{M}^{\psi}_{Ax} \Leftrightarrow s_i \notin S_i^{\psi} \Leftrightarrow s_i \in \wp(OL) \backslash S_i^{\psi} = S_i^{\phi} \Leftrightarrow M \in \mathcal{M}^{\phi}_{Ax}$.

$\phi = \psi_1 \wedge \psi_2$: By the induction hypothesis, $M \models \psi_1 \Leftrightarrow M \in \mathcal{M}^{\psi_1}_{Ax}$ and $M \models \psi_2 \Leftrightarrow M \in \mathcal{M}^{\psi_2}_{Ax}$. $M \models \phi \Leftrightarrow M \models \psi_1$ and $M \models \psi_2 \Leftrightarrow M \in \mathcal{M}^{\psi_1}_{Ax}$ and $M \in \mathcal{M}^{\psi_2}_{Ax} \Leftrightarrow s_i \in S_i^{\psi_1}$ and $s_i \in S_i^{\psi_2} \Leftrightarrow s_i \in S_i^{\psi_1} \cap S_i^{\psi_2} = S_i^{\phi} \Leftrightarrow M \in \mathcal{M}^{\phi}_{Ax}$. $\blacksquare$

Of course, the model class $\mathcal{M}^{\Phi}_{Ax}$ of a set of formulae $\Phi \subseteq Ax$ (see Eq. 6.1 on p. 31) can be constructed by restricting the set of allowed epistemic states locally for each agent. Let S_i^{ϕ} be as defined in Def. 6.4, and let

$$S_i^{\Phi} = \bigcap_{\phi \in (\Phi \cap EF^i)} S_i^{\phi}$$

Then,

$$\mathcal{M}^{\Phi}_{Ax} = \{(s_1^{\Phi}, \ldots, s_n^{\Phi}, \pi) : s_i^{\Phi} \in S_i^{\Phi}\} \tag{6.2}$$

6.3 Common Axioms

Classical modal epistemic logic, henceforth "epistemic logic", defines the basic concept of knowledge (however we interpret this concept). In addition, axioms corresponding to properties of the agents are often added to the basic system in order to create derived systems. There is a whole catalogue of such systems, corresponding to the different combinations of (additional) axioms. The most common axioms from epistemic logic, written in our notation, are ($1 \leq i \leq n$):

$\triangle_i \{(\alpha \to \beta)\} \to (\triangle_i\{\alpha\} \to \triangle_i\{\beta\})$	Distribution	$\mathbf{K_i}$
$\triangle_i \{\alpha\} \to \neg \triangle_i \{\neg\alpha\}$	Consistency	$\mathbf{D_i}$
$\triangle_i \{\alpha\} \to \alpha$	Knowledge	$\mathbf{T_i}$
$\triangle_i \{\alpha\} \to \triangle_i\{\triangle_i\{\alpha\}\}$	Positive Introspection	$\mathbf{4_i}$
$\neg \triangle_i \{\alpha\} \to \triangle_i\{\neg \triangle_i \{\alpha\}\}$	Negative Introspection	$\mathbf{5_i}$

A logical system is *characterized* by a class of models if and only if it is sound and complete with respect to that class. An important result in epistemic logic (and in modal logic in general) is the *correspondence* between certain axioms and model classes. For example, the basic system only including $\mathbf{K_i}$ of the axioms above is characterized by the class of all (multi-modal) Kripke structures, and the system only including $\mathbf{K_i}$ and $\mathbf{T_i}$ of the axioms above is characterized by the class of all reflexive Kripke structures.

In order to compare our framework with the traditional one, we can now show similar characterizations by extending our logical system with axioms. By adding a set of epistemic axioms Φ to EC we obtain a logical system which we call EC_Φ, and the corresponding inference relation is denoted $\vdash_\Phi$. It is easy to see that the system EC_Φ is sound and complete with respect to the structure class $\mathcal{M}^\Phi$, i.e. that

$$\Gamma \vdash_\Phi \phi \Leftrightarrow \Gamma \models_{\mathcal{M}^\Phi} \phi$$

because $\Gamma \models_{\mathcal{M}^\Phi} \phi$ iff $\Phi \cup \Gamma \models \phi$ iff — by soundness and completeness of EC — $\Phi \cup \Gamma \vdash \phi$ iff — because the proof is identical whether we use Φ as axioms or as premises — $\Gamma \vdash_\Phi \phi$. In other words; when we add a set of axioms $\Phi \subseteq Ax$ to EC, the resulting system is characterized by the class of structures $\mathcal{M}^\Phi$ (which we showed how to construct in the previous section). This structure class has a simple description: it is constructed by removing certain epistemic states from the set of possible epistemic states. For example, the system we obtain by including $\mathbf{D_i}$ is characterized by the set of all structures which do not contain both a formula and its negation in an epistemic state.

We construct the model classes for the axioms, except for $\mathbf{T_i}$ which is not an epistemic axiom (as defined in the previous section). We interpret the axioms as schemata over both the agent index i and the agent language formulae α, β. In order to construct the model classes, we must view these axiom schemata as sets of axioms:

$$\begin{array}{ll}
\mathbf{K} = \{K_i(\alpha,\beta) : 1 \leq i \leq n, \alpha,\beta \in AL\} & \text{where } K_i(\alpha,\beta) = \\
& \triangle_i \{(\alpha \to \beta)\} \to (\triangle_i\{\alpha\} \to \triangle_i\{\beta\}) \\
\mathbf{D} = \{D_i(\alpha) : 1 \leq i \leq n, \alpha \in AL\} & \text{where } D_i(\alpha) = \triangle_i\{\alpha\} \to \neg \triangle_i \{\neg\alpha\} \\
\mathbf{4} = \{4_i(\alpha) : 1 \leq i \leq n, \alpha \in AL\} & \text{where } 4_i(\alpha) = \triangle_i\{\alpha\} \to \triangle_i\{\triangle_i\{\alpha\}\} \\
\mathbf{5} = \{5_i(\alpha) : 1 \leq i \leq n, \alpha \in AL\} & \text{where } 5_i(\alpha) = \neg \triangle_i \{\alpha\} \to \triangle_i\{\neg \triangle_i \{\alpha\}\}
\end{array}$$

The model classes are defined by eq. 6.2, where:

$$\begin{array}{ll}
S_i^{K_i(\alpha,\beta)} & = \wp(OL) \setminus \{X \in \wp(OL) : [\alpha \to \beta] \in X\} \cup (\wp(OL) \setminus \{X \in \wp(OL) : [\alpha] \in X\} \\
& \cup \{X \in \wp(OL) : [\beta] \in X\}) \\
& = \wp(OL) \setminus \{X \in \wp(OL) : ([\alpha] \to [\beta]), [\alpha] \in X, [\beta] \notin X\} \\
S_i^{D_i(\alpha)} & = \wp(OL) \setminus \{X \in \wp(OL) : [\alpha] \in X\} \cup \wp(OL) \setminus \{X \in \wp(OL) : \neg[\alpha] \in X\} \\
& = \wp(OL) \setminus \{X \in \wp(OL) : [\alpha], \neg[\alpha] \in X\} \\
S_i^{4_i(\alpha)} & = \wp(OL) \setminus \{X \in \wp(OL) : [\alpha] \in X\} \cup \{X \in \wp(OL) : \triangle_i\{[\alpha]\} \in X\} \\
& = \wp(OL) \setminus \{X \in \wp(OL) : [\alpha] \in X, \triangle_i\{[\alpha]\} \notin X\} \\
S_i^{5_i(\alpha)} & = \wp(OL) \setminus (\wp(OL) \setminus \{X \in \wp(OL) : [\alpha] \in X\}) \cup \{X \in \wp(OL) : \neg \triangle_i \{[\alpha]\} \in X\} \\
& = \wp(OL) \setminus \{X \in \wp(OL) : [\alpha] \notin X, \neg \triangle_i \{[\alpha]\} \notin X\} \\
S_i^{\mathbf{K}} & = \bigcap_{K_i(\alpha,\beta) \in \mathbf{K}} S_i^{K_i(\alpha,\beta)} = \\
& \wp(OL) \setminus \{X \in \wp(OL) : \exists \alpha,\beta \in AL(([\alpha] \to [\beta]), [\alpha] \in X, [\beta] \notin X\}) \\
S_i^{\mathbf{D}} & = \bigcap_{D_i(\alpha) \in \mathbf{D}} S_i^{D_i(\alpha)} = \wp(OL) \setminus \{X \in \wp(OL) : \exists \alpha \in AL([\alpha], \neg[\alpha] \in X\}) \\
S_i^{\mathbf{4}} & = \bigcap_{4_i(\alpha) \in \mathbf{4}} S_i^{4_i(\alpha)} = \wp(OL) \setminus \{X \in \wp(OL) : \exists \alpha \in AL([\alpha] \in X, \triangle_i\{[\alpha]\} \notin X\}) \\
S_i^{\mathbf{5}} & = \bigcap_{5_i(\alpha) \in \mathbf{5}} S_i^{5_i(\alpha)} = \wp(OL) \setminus \{X \in \wp(OL) : \exists \alpha \in AL([\alpha] \notin X, \neg \triangle_i \{[\alpha]\} \notin X\})
\end{array}$$

Axiom	Kripke property	Illegal KSS-state s	Non-triv. sat. in $\mathcal{M}_{fin}$?
$\mathbf{K_i}$	none	$[\alpha], ([\alpha] \rightarrow [\beta]) \in s, [\beta] \notin s$	yes
$\mathbf{D_i}$	serial	$[\alpha], \neg[\alpha] \in s$	yes
$\mathbf{T_i}$	refl.	*(Not an epistemic axiom)*	
$\mathbf{4_i}$	trans.	$[\alpha] \in s, \triangle_i\{[\alpha]\} \notin s$	no
$\mathbf{5_i}$	Eucl.	$[\alpha] \notin s, \neg \triangle_i \{[\alpha]\} \notin s$	no

Table 6.1: Summary of properties of axioms and related systems. The second column shows the property of the class of Kripke structures which characterizes the system including the axiom in the first column. The third column shows the property of the class of Knowledge Set structures which characterizes the system; i.e. the epistemic states which are "illegal" in a structure (in the notation from Definition 6.4, the column shows the property of s such that $s \notin S_i^{\phi}$). The fourth column says whether the axiom is *non-trivially* satisfiable in $\mathcal{M}_{fin}$, i.e. whether it can be satisfied by models in $\mathcal{M}_{fin}$ other than the trivial models where agent i has an empty epistemic state ($s_i = \emptyset$).

Table 6.1 shows a summary of characterizations of the different axioms. The model class $mod(\mathbf{T_i})$ cannot be described by a local property of the set of allowed epistemic states.

An interesting property of an epistemic axiom is whether it can be used to describe finite agents (see Chapter 5). As shown (Tab. 6.1), only $\mathbf{K_i}$ and $\mathbf{D_i}$ fall into this category. The only models in $\mathcal{M}_{fin}$ satisfying $\mathbf{4_i}$ are the uninteresting models where agent i does not know anything, so if we adopt $\mathbf{4_i}$ an agent who knows at least one fact must know infinitely many facts. $\mathbf{5_i}$ are not even satisfied by trivial models, so if we adopt $\mathbf{5_i}$ every agent must know infinitely many facts. This observation points out an interesting dichotomy between potential epistemic axioms with respect to the logical omniscience problem. $\mathbf{4_i}$ and $\mathbf{5_i}$ imply a form of logical omniscience, while $\mathbf{K_i}$ and $\mathbf{D_i}$ can be used to describe the fact that there are certain epistemic states which the agents never go into in their deliberation process.

6.4 On An Incompleteness of Knowledge

We show a property of our notion of explicitly only knowing. In the following theorem, we adopt the $\mathbf{T_i}$ axioms (knowledge is truth).

Theorem 6.6 *For any term S:*

$$\mathbf{T_i} \models \neg \triangle_i \{\triangledown_i S\} \qquad \blacksquare$$

PROOF It is easy to see that $\mathbf{T_i}$ is satisfiable[1]. Let $M \models \mathbf{T_i}$, and assume that $M \models \triangle_i\{\triangledown_i S\}$ for some S, i.e. that $\triangledown_i[S] \in s_i$. By $\mathbf{T_i}$, $M \models \triangledown_i S$ and $s_i \subseteq [S]$. Then $\triangledown_i[S] \in [S]$, which is impossible. $\blacksquare$

Theorem 6.6 can be seen as an epistemic analogue of Gödels incompleteness theorem: there exist true formulae which are impossible to know — more specifically, it is impossible to know that "this is all I know" (if knowledge is truth). Furthermore, this holds even for (imaginary) agents with infinite memory[2].

[1] Consider for example a structure with $s_i = \{p\}$ and $\pi(p) = true$.

[2] A formula like $\triangle_i\{\triangledown_j S\}$, $i \neq j$, is of course fully consistent with $\mathbf{T_i}$; an agent can potentially know all *another* agent knows.

Example 6.7: *Consider once again the encrypted message example from the introduction. We can express the fact that agent a knows that agent b knows at most* m_e *and* n *by the formula*

$$\phi = \triangle_a\{\triangledown_b\{m_e, n\}\}$$

It is easy to see that ϕ *is satisfiable in* $mod(\mathbf{T_i})$*; let* $\triangledown_b\{m_e, n\} \in s_a$ *and* $s_b \subseteq \{m_e, n\}$.

The fact that agent b knows that it knows at most m_e *and* n*, i.e.*

$$\phi_1 = \triangle_b\{\triangledown_b\{m_e, n\}\}$$

is clearly unsatisfiable in $\mathcal{M}^{\mathbf{T_i}}$*. Since the epistemic state of b must include the formula* $\triangledown_b\{m_e, n\}$ *but in* $\mathcal{M}^{\mathbf{T_i}}$ *in every model of* ϕ_1 *the formula*

$$\triangledown_b\{m_e, n\}$$

must be true, i.e. the epistemic state of b can at most include m_e *and* n*. If we add this formula to the description of the epistemic state of agent b, we get*

$$\phi_2 = \triangle_b\{\triangledown_b\{m_e, n, \triangledown_b\{m_e, n\}\}\}$$

but in every model of ϕ_2 *the formula*

$$\triangledown_b\{m_e, n, \triangledown_b\{m_e, n\}\}$$

must be true, so ϕ_2 *is also unsatisfiable. Continuing in the same way, we get*

$$\phi_3 = \triangle_b\{\triangledown_b\{m_e, n, \triangledown_b\{m_e, n\}, \triangledown_b\{m_e, n, \triangledown_b\{m_e, n\}\}\}\}$$
$$\vdots$$

ad infinitum, and there is no ϕ_i *which is satisfiable in* $\mathcal{M}^{\mathbf{T_i}}$. □

The property (Th. 6.6) is a consequence of an impossibility of self-reference of terms: a (finite) term cannot be its own proper subterm.

Chapter 7

Related Work, Conclusions and Further Work

The idea that an agent knows — whatever the interpretation of this concept — a set of formulae at a given instant is implicitly assumed in most approaches to reasoning about knowledge (the set known by agent i in model M is $\{\phi : M \models K_i\phi\}$). Semantically, some approaches use this assumption explicitly in the definition of knowledge rather than defining knowledge as truth in all possible worlds (Eberle, 1974; Moore & Hendrix, 1979; Halpern & Moses, 1990; Scott, 1970; Konolige, 1986a). Our logic EL, in addition, takes the assumption one step further and includes explicit reasoning about sets of formulae and the agents' (possibly nested) knowledge of them in the logical language. Our approach uses a *meta-language* in which the (sets of) agent language formulae is represented as terms, as opposed to modal languages where the argument to the knowledge operator is a proposition.

7.1 Related Work

The main idea in our approach is to combine the ideas of explicit knowledge and only knowing. We review some other approaches to these two separate concepts.

7.1.1 Explicit Knowledge

Since Levesque (1984b) (revised in (Levesque, 1984a)) defined explicit knowledge as the *actual* knowledge held explicitly by an agent and implicit knowledge as that which follow logically from explicit knowledge, formalizations of these concepts have been a recurrent theme in the literature. The main reason for this is the logical omniscience problem; an agent can only act upon what he explicitly knows. In fact, most suggested solutions to the logical omniscience problem can be seen as models of explicit knowledge. For example, the knowledge computed by a resource-bounded agent is explicit knowledge.

Many of the approaches set out to find rules governing explicit knowledge. One of them is Levesque's seminal paper (Levesque, 1984b). Levesque notes that a solution to the logical omniscience problem would be to interpret knowledge in standard modal epistemic logic as implicit knowledge. Of course, this only changes the problem to finding a logic that describes explicit knowledge. He introduces a language[1] with two modalities — for implicit and explicit knowledge — and proposes a semantics based on impossible worlds. It turns out that Levesque's

[1]The language Levesque considers is quite restricted. Particularly, nested knowledge is not allowed. The framework has later been extended to allow nested knowledge (Lakemeyer, 1987).

explicit knowledge can be characterized by the notion of entailment in relevance logic (Anderson & Belnap, 1975), and the corresponding axioms and rules can thus be "imported" directly from relevance logic. Some of them are[2]:

$$B(\phi \wedge \psi) \leftrightarrow B(\psi \wedge \phi) \tag{7.1}$$

$$B\neg(\phi \vee \psi) \leftrightarrow B(\neg\phi \wedge \neg\psi) \tag{7.2}$$

$$B\phi \wedge B\psi \leftrightarrow B(\phi \wedge \psi) \tag{7.3}$$

$$\frac{\vdash ((B\phi \vee B\psi) \to B\gamma)}{\vdash B(\phi \vee \psi) \to B\gamma} \tag{7.4}$$

These axioms and rules illustrate that Levesque's notion of explicit knowledge is closed under "simple" consequences.

One approach to modeling the concepts of implicit and explicit knowledge which does not set out to model rules governing explicit knowledge is the *logic of General Awareness* (Fagin & Halpern, 1988). The main idea is that in order to say that an agent knows a fact ϕ, 1) ϕ must follow from the agent's information and 2) the agent must be *aware* of ϕ. Three different unary operators are used in the language to account for the concepts of implicit knowledge, explicit knowledge and awareness. An agent knows a fact implicitly if it is true in all possible worlds and knows a fact explicitly if he knows it implicitly and he is aware of it. Awareness of a fact in a situation is given by a syntactical awareness function, assigning a set of formulae the agent is aware of in each situation. Fagin and Halpern do not attach any fixed meaning to this notion of awareness, but leave it open to interpretation. However, it can be argued that the suggested possible interpretations of awareness do not make the definition of explicit knowledge very plausible (Hadley, 1986, 1988; Konolige, 1986b). The idea of using awareness to model explicit knowledge has also been used by others (Huang & Kwast, 1991; Thijsse, 1993).

7.1.2 Only Knowing

Several authors have analyzed the knowledge state of an agent who knows a (set of) formula(e) (Konolige, 1982; Moore, 1983; Halpern & Moses, 1985; Halpern, 1997). Levesque (1990) introduced a logic in which *only knowing* can be expressed inside the logical language. Briefly speaking, Levesque's language is of first order[3] and has two unary epistemic connectives **B** and **O**.[4] Semantically, a *world* is a truth assignment to the primitive sentences, and satisfaction of a formula is defined relative to a pair W, w where W is the set of worlds the agent considers possible and w is the "real" world[5] (the world corresponding to the correct state of affairs). A sentence $\mathbf{B}\alpha$ is true in W, w iff α is true in W, w' for every $w' \in W$; **B** is the traditional belief/knowledge operator in modal epistemic logic. A sentence $\mathbf{O}\alpha$ is true in W, w iff $\mathbf{B}\alpha$ is true in W, w and $w' \in W$ for every w' such that α is true in W, w'. $\mathbf{O}\alpha$ expresses that the agent only knows α; the set of possible worlds is as large as possible consistent with believing α. It is shown that the **O** operator can be modeled by a "natural dual" to the **B** operator — an operator **N**. The intended meaning of $\mathbf{N}\alpha$ is that α at most is believed to be false, and $\mathbf{N}\alpha$ is true in W, w iff $w' \in W$ for every w' such that α is false in W, w'. Then, $\mathbf{O}\alpha$ is true iff $\mathbf{B}\alpha$ and $\mathbf{N}\neg\alpha$ is true; **B** specifies a lower bound and **N** specifies an upper bound on what is believed.

[2] The last rule was erroneously omitted from (Levesque, 1984b), but was included in (Levesque, 1984a).

[3] The logic was only shown to be complete for the unquantified version of the language, the full version was later shown to be incomplete (Halpern & Lakemeyer, 1995).

[4] Levesque only considers a single agent, but his approach has later been extended to the multi-agent case (Halpern & Lakemeyer, 1996)

[5] Note that this corresponds to the semantical assumptions of the modal logic $S5$ for one agent.

Although the concept of only knowing obviously is useful, Levesque's approach suffers from the logical omniscience problem. The **O** operator tells which equivalence class of formulae the agent knows; knowledge is closed under logical equivalence — a clearly unrealistic assumption.

7.2 Conclusions

A proper logic for reasoning about knowledge must combine the concepts of explicit knowledge and only knowing. We have presented a logic which does this by allowing an explicit representation of sets of formulae in the logical language. The main semantical idea is simple and natural: an agent is represented as a point in the subset lattice of all formulae. As opposed to taking knowledge as a propositional attitude, this semantical approach is mainly syntactical; the only assumption about equivalent formula with respect to knowledge is that the agents can represent *sets*, e.g. that if an agent knows $\{\alpha, \beta\}$ he also knows $\{\beta, \alpha\}$ and $\{\alpha, \alpha, \beta\}$. The logical language we have constructed is a meta-language, as opposed to a modal type language.

Our logic, as opposed to Levesque's (Levesque, 1984b), is not an attempt to describe the rules governing explicit knowledge. Rather, the syntactical semantic approach makes no assumption about which epistemic states are legal, so there is no logic describing the content of these states (in general; we will discuss extensions of the approach shortly). A critique of "[characterizing] a state of knowledge syntactically, by the set of formulas that are true in this state" is offered by Fagin, Halpern, & Vardi (1991, p. 3): "Describing a state by the formulas that are true in it seems to avoid the issue of modeling altogether". This is a valid critique of approaches which (only) model the traditional knowledge concept. Our approach, however, in addition model interesting concepts for reasoning about knowledge states and is thus not merely a logic for explicit knowledge *representation*.

It is clear that an infinite storage for explicit knowledge is not much more plausible than logical omniscience. We model finite agents semantically by considering only finite subsets of all formulae. We cannot, however, reflect this in the logical system by using finite axioms or rules; certain infinite sets of formulae describe infinite agents but have no finite subsets which does the same. Completeness for finite agents is restricted to *finitary* theories. Particularly, the set of theorems is equal to the set of valid formulae (with respect to the semantics for finite agents), and if Γ is a finitary (e.g. a finite) theory then a formula is provable from the premises Γ if and only if it is a logical consequence of Γ.

Our approach is based on simple epistemic assumptions. We can also model additional epistemic properties of the agents. We have shown that purely epistemic axioms corresponds to removing epistemic states from the set of legal epistemic states for each agent individually, and we have given a constructive definition of the set of legal states.

Although our meta language contains terms which are formulae in the meta language itself, it is not self-referential in the sense that an expression can refer to itself (the definition of a term may not involve infinite recursion). This is the reason that, if knowledge implies truth, it is impossible to know "all I know" or, in other words, an agent can never know any true upper bound on its knowledge.

7.3 Future Work

We have modeled concepts related to static sets of explicit knowledge. Currently, we are working on extending this framework with a model of the dynamics of how knowledge becomes explicit, in the spirit of (Konolige, 1986a).

An interesting aspect of "only knowing" is only knowing *about* something. We were able to express the fact that agent b in the encrypted message example from the introduction only knows m_e and n by the formula $\Diamond_b\{m_e, n\}$. However, agent b might also know several other facts not related to the problem in question (the encrypted message), such as the fact that it is raining. It clearly would be useful to extend the logic in order to be able to express all agent b knows about the encrypted message only, instead of having to list all he knows about anything. This would involve a model of *relevance* for determining which formulae is "about" e.g. the encrypted message.

An extension to a first-order language would also be interesting.

Bibliography

Anderson, A. R., and Belnap, N. D. J. 1975. *Entailment, the Logic of Relevance and Necessity*. Princeton: Princeton University Press.

Eberle, R. A. 1974. A logic of believing, knowing and inferring. *Synthese* 26:356–382.

Fagin, R., and Halpern, J. Y. 1985. Belief, awareness and limited reasoning. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, 491–501.

Fagin, R., and Halpern, J. Y. 1988. Belief, awareness and limited reasoning. *Artificial Intelligence* 34:39–76. A preliminary version appeared in Fagin & Halpern (1985).

Fagin, R.; Halpern, J. Y.; and Vardi, M. Y. 1991. A model-theoretic analysis of knowledge. *Journal of the ACM* 38(2):382–428. A preliminary version appeared in *Proc. 25th IEEE Symposium on Foundations of Computer Science*, 1984.

Hadley, R. F. 1986. Fagin and halpern on logical omniscience: A critique with an alternative. In *Proc. Sixth Canadian Conference on Artificial Intelligence*, 49 – 56. Montreal: University of Quebec Press.

Hadley, R. F. 1988. Logical omniscience, semantics and models of belief. *Computational Intelligence* 4:17–30.

Halpern, J. Y., and Lakemeyer, G. 1995. Levesque's axiomatization of only knowing is incomplete. *Artificial Intelligence* 74(2):381–387.

Halpern, J. Y., and Lakemeyer, G. 1996. Multi-agent only knowing. In Shoham, Y., ed., *Theoretical Aspects of Rationality and Knowledge: Proceedings of the Sixth Conference (TARK 1996)*. San Francisco: Morgan Kaufmann. 251–265.

Halpern, J. Y., and Moses, Y. 1985. Towards a theory of knowledge and ignorance. In Apt, K. R., ed., *Logics and Models of Concurrent Systems*. Berlin: Springer-Verlag. 459–476.

Halpern, J. Y., and Moses, Y. 1990. Knowledge and common knowledge in a distributed environment. *Journal of the ACM* 37(3):549–587.

Halpern, J. Y. 1997. A theory of knowledge and ignorance for many agents. *Journal of Logic and Computation* 7(1):79–108.

Huang, Z., and Kwast, K. 1991. Awareness, negation and logical omniscience. In van Eijk, J., ed., *Logics in AI, Proceedings JELIA'90*, volume 478 of *Lecture Notes in Computer Science*. Berlin: Springer-Verlag. 282–300.

Konolige, K. 1982. Circumscriptive ignorance. In Waltz, D., ed., *Proceedings of the National Conference on Artificial Intelligence*, 202–204. Pittsburgh, PA: AAAI Press.

Konolige, K. 1986a. *A Deduction Model of Belief.* Los Altos, California: Morgan Kaufmann Publishers.

Konolige, K. 1986b. What awareness isn't: A sentential view of implicit and explicit belief. In Halpern, J. Y., ed., *Theoretical Aspects of Reasoning About Knowledge: Proceedings of the First Conference*, 241–250. Los Altos, California: Morgan Kaufmann Publishers, Inc.

Lakemeyer, G. 1987. Tractable meta-reasoning in propositional logics of belief. In McDermott, J., ed., *Proceedings of the 10th International Joint Conference on Artificial Intelligence (IJCAI '87)*, 401–408. Milan, Italy: Morgan Kaufmann.

Levesque, H. J. 1984a. A logic of implicit and explicit belief. Technical Report 32, Fairchild Laboratory of Artificial Intelligence, Palo Alto, US.

Levesque, H. J. 1984b. A logic of implicit and explicit belief. In *Proceedings of the Fourth National Conference on Artificial Intelligence (AAAI'84)*, 198–202. American Association for Artificial Intelligence.

Levesque, H. J. 1990. All I know: a study in autoepistemic logic. *Artificial Intelligence* 42:263–309.

Moore, R. C., and Hendrix, G. 1979. Computational models of beliefs and the semantics of belief sentences. Technical Note 187, SRI International, Menlo Park, CA.

Moore, R. C. 1983. Semantical considerations on nonmonotonic logic. In Bundy, A., ed., *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, 272–279. Karlsruhe, FRG: William Kaufmann.

Scott, D. S. 1970. Advice on modal logic. In Lambert, K., ed., *Philosophical Problems in Logic*. Dordrecht: D. Reidel Publishing Co. 143–173.

Thijsse, E. 1993. On total awareness logics. In de Rijke, M., ed., *Diamonds and Defaults*. Dordrecht: Kluwer Academic Publishers. 309–347.

Appendix A

A Fixed-point Characterization of *EL*

The definitions of AL (Def. 3.4), $TL(AL)$ (Def. 3.1) and EL (Def. 3.5) are collected in Definition A.1 below.

In the following we write just TL for $TL(AL)$. Also, we shall be pedantic with respect to the notation for basic terms and use the underlined notation.

The three languages are defined over the alphabet

$$A = \Theta \cup \{\triangle_i, \triangledown_i, \neg, \wedge, (,), \doteq \underline{\{}, \underline{\}}, \underline{,}\} \tag{A.1}$$

We use the following notation for the set of descriptors for finite subsets of a set $X \subseteq A^*$:

$$Fin(X) = \{\underline{\{}\alpha_1\underline{,} \ldots \underline{,}\alpha_k\underline{\}} : \alpha_i \in X, k \geq 1\} \tag{A.2}$$

Definition A.1:

- *TL is the least set such that*
  - $Fin(AL) \subseteq TL$
  - *If* $T, U \in TL$ *then* $\left.\begin{array}{c}(T \sqcup U) \\ (T \sqcap U)\end{array}\right\} \in TL$
- *AL is the least set such that*
  - $\Theta \subseteq AL$
  - *If* $T \in TL$ *then* $\left.\begin{array}{c}\triangle_i T \\ \triangledown_i T\end{array}\right\} \in AL$ *(* $1 \leq i \leq n$ *)*
  - *If* $\alpha, \beta \in AL$ *then* $\left.\begin{array}{c}\neg\alpha \\ (\alpha \wedge \beta)\end{array}\right\} \in AL$
- *EL is the least set such that*
  - $AL \subseteq EL$
  - *If* $T, U \in TL$ *then* $(T \doteq U) \in EL$
  - *If* $\phi, \psi \in EL$ *then* $\left.\begin{array}{c}\neg\phi \\ (\phi \wedge \psi)\end{array}\right\} \in EL$ □

We show that the three sets are well-defined.

Lemma A.2 *The recursive Definition A.1 has a solution.* ■

PROOF We prove that the associated generating function has a fixed point.

To make the discussion easier, we rewrite Definition A.1 as a set of three mutually recursive equations:

$$TL = Fin(AL) \cup \{T \sqcup U : T, U \in TL\} \cup \{T \sqcap U : T, U \in TL\} \tag{A.3}$$

$$\begin{aligned} AL =& \Theta \cup \{\triangle_i T : T \in TL\} \cup \{\triangledown_i T : T \in TL\} \cup \\ & \{\neg\phi : \phi \in AL\} \cup \{(\phi \wedge \psi) : \phi, \psi \in AL\} \end{aligned} \tag{A.4}$$

$$\begin{aligned} EL =& AL \cup \{T \doteq U : T, U \in TL\} \cup \\ & \{\neg\phi : \phi \in EL\} \cup \{(\phi \wedge \psi) : \phi, \psi \in EL\} \end{aligned} \tag{A.5}$$

We define a function

$$g : \wp(A^*)^3 \to \wp(A^*)^3 \tag{A.6}$$

as a generating function corresponding to the recursive definitions, as follows (the names of the variables have been changed):

$$g(TL', AL', EL') = \langle g_T(TL', AL', EL'), g_O(TL', AL', EL'), g_E(TL', AL', EL')\rangle \tag{A.7}$$

where

$$g_T(TL', AL', EL') = Fin(AL') \cup \{T \sqcup U : T, U \in TL'\} \cup \{T \sqcap U : T, U \in TL'\} \tag{A.8}$$

$$\begin{aligned} g_O(TL', AL', EL') =& \Theta \cup \{\triangle_i T : T \in TL'\} \cup \{\triangledown_i T : T \in TL'\} \cup \\ & \{\neg\phi : \phi \in AL'\} \cup \{(\phi \wedge \psi) : \phi, \psi \in AL'\} \end{aligned} \tag{A.9}$$

$$\begin{aligned} g_E(TL', AL', EL') =& AL' \cup \{T \doteq U : T, U \in TL'\} \cup \\ & \{\neg\phi : \phi \in EL'\} \cup \{(\phi \wedge \psi) : \phi, \psi \in EL'\} \end{aligned} \tag{A.10}$$

Then, we define an ordering relation $\sqsubseteq$ on $\wp(A^*)^3$:

$$\langle T, O, E\rangle \sqsubseteq \langle T', O', E'\rangle \Leftrightarrow \begin{cases} T \subseteq T' & \text{and} \\ O \subseteq O' & \text{and} \\ E \subseteq E' \end{cases} \tag{A.11}$$

$\sqsubseteq$ is a pointed (has bottom) complete partial order, so if we can prove that g is continuous with respect to $\sqsubseteq$ we have proved that it has a least fixed point.

Define the least upper bound $c_1 \sqcup_{\sqsubset} c_2$ of two elements $c_1, c_2 \in \wp(A^*)^3$ in the usual way. It is easy to see that, for $S \subseteq \wp(A^*)^3$

$$\bigsqcup_{\sqsubset} S = \langle \bigcup\{T : \langle T, O, E\rangle \in S\}, \bigcup\{O : \langle T, O, E\rangle \in S\}, \bigcup\{E : \langle T, O, E\rangle \in S\}\rangle \tag{A.12}$$

g is continuous iff for every chain C in $\wp(A^*)^3$,

$$g(\bigsqcup_{\sqsubset} C) = \bigsqcup_{\sqsubset}\{g(c) : c \in C\} \tag{A.13}$$

Henceforth, let $C = \{c_1, c_2, \ldots\}$ where $c_i \sqsubseteq c_{i+1}$ be an arbitrary chain in $\wp(A^*)^3$[1], and let $c_i = \langle T_i, O_i, E_i\rangle$. Clearly, $\{T_i : c_i \in C\}$, $\{O_i : c_i \in C\}$ and $\{E_i : c_i \in C\}$ are all chains in $\wp(A^*)$, and $T_i \subseteq T_{i+1}$, $O_i \subseteq O_{i+1}$ and $E_i \subseteq E_{i+1}$.

[1] A chain in $\wp(A^*)^3$ is countable.

Let $S = \{S_1, S_2, \ldots\}$ be a chain in $\wp(A^*)$. We show two properties of S. The first is:

$$Fin(\bigcup_i \{S_i\}) = \bigcup_i \{Fin(S_i)\} \tag{A.14}$$

We show that this equation holds. First, let $x \in \bigcup_i \{Fin(S_i)\}$. Then, $x \in Fin(S_j)$ for some j, and $x = \{\alpha_1, \ldots, \alpha_k\}$ where $\alpha_i \in S_j$. Since $S_j \subseteq \bigcup_i \{S_i\}$, $\alpha_i \in \bigcup_i \{S_i\}$ and thus $x \in Fin(\bigcup_i \{S_i\})$. Second, let $x = \{\alpha_1, \ldots, \alpha_k\} \in Fin(\bigcup_i \{S_i\})$. Then $\alpha_i \in \bigcup_i \{S_i\}$, so for each α_i there is a S'_i such that $\alpha \in S'_i$.

The chain $S'_1, \ldots, S'_k$ is finite, so it contains it's least upper bound – let the lub be S'_j. Since $S'_1 \cup \ldots \cup S'_k = S'_j$, we have that $\alpha_i \subseteq S'_j$ for all i and so $x \in Fin(S'_j)$. Since $S'_j \in S$, $x \in \bigcup_i \{Fin(S_i)\}$.

The second property of the chain S we need is:

$$x_1, x_2 \in \bigcup_i \{S_i\} \text{ iff there is some } S_j \text{ such that } x_1, x_2 \in S_j \tag{A.15}$$

If $x_1, x_2 \in \bigcup_i \{S_i\}$ then there exist S_1, S_2 such that $x_1 \in S_1$ and $x_2 \in S_2$. If $S_1 \subseteq S_2$ then $x_1, x_2 \in S_2$ and if $S_2 \subseteq S_1$ then $x_1, x_2 \in S_1$. Conversely, if $x_1, x_2 \in S_j$ for some S_j then $x_1, x_2 \in \bigcup_i \{S_i\}$ trivially.

In order to show eq. A.13, we first show that

$$g_T(\bigsqcup_{\sqsubset} C) = \bigcup \{g_T(c) : c \in C\} \tag{A.16}$$

$$\begin{aligned}
g_T(\bigsqcup_{\sqsubset} C) &= g_T(\bigcup_i \{T_i\}, \bigcup_i \{O_i\}, \bigcup_i \{E_i\}) && \text{By (A.12)} && \text{(A.17)}\\
&= Fin(\bigcup_i \{O_i\}) \cup \{T \sqcup U : T, U \in \bigcup_i \{T_i\}\} \\
&\quad \cup \{T \sqcap U : T, U \in \bigcup_i \{T_i\}\} && && \text{(A.18)}\\
&= \bigcup_i \{Fin(O_i)\} \cup \{T \sqcup U : T, U \in \bigcup_i \{T_i\}\} \\
&\quad \cup \{T \sqcap U : T, U \in \bigcup_i \{T_i\}\} && \text{By (A.14)} && \text{(A.19)}\\
&= \bigcup_i \{Fin(O_i)\} \cup \{T \sqcup U : T, U \in T_j \text{for some } T_j\} \\
&\quad \cup \{T \sqcap U : T, U \in T_j \text{for some } T_j\} && \text{By (A.15)} && \text{(A.20)}\\
&= \bigcup_i \{Fin(O_i)\} \cup \bigcup_i \{T \sqcup U : T, U \in T_i\} && && \text{(A.21)}\\
&\quad \cup \bigcup_i \{T \sqcap U : T, U \in T_i\} && && \text{(A.22)}\\
&= \bigcup \{Fin(O_i) \cup \{T \sqcup U : T, U \in T_i\} \\
&\quad \cup \{T \sqcap U : T, U \in T_i\} : \langle T_i, O_i, E_i \rangle \in C\} && && \text{(A.23)}\\
&= \bigcup \{g_T(c) : c \in C\} && && \text{(A.24)}
\end{aligned}$$

Second, we show that

$$g_O(\bigsqcup_{\sqsubset} C) = \bigcup \{g_O(c) : c \in C\} \tag{A.25}$$

$$
\begin{aligned}
g_O(\bigsqcup_{\sqsubset} C) &= g_O(\bigcup_i\{T_i\},\bigcup_i\{O_i\},\bigcup_i\{E_i\}) && \text{By (A.12)}\\
&=\Theta \cup \{\triangle_i T : T \in \bigcup_i\{T_i\}\} \cup \{\triangledown_i T : T \in \bigcup_i\{T_i\}\}\cup \\
&\quad \{\neg\phi : \phi \in \bigcup_i\{O_i\}\} \cup \{(\phi\wedge\psi) : \phi,\psi \in \bigcup_i\{O_i\}\} && \text{(A.26)}\\
&=\Theta \cup \{\triangle_i T : T \in \bigcup_i\{T_i\}\} \cup \{\triangledown_i T : T \in \bigcup_i\{T_i\}\}\cup \\
&\quad \{\neg\phi : \phi \in \bigcup_i\{O_i\}\} \cup \{(\phi\wedge\psi) : \phi,\psi \in O_j \text{for some } O_j\} && \text{By (A.15)}\\
&=\Theta \cup \bigcup_i\{\triangle_i T : T \in T_i\} \cup \bigcup_i\{\triangledown_i T : T \in T_i\}\cup \\
&\quad \bigcup_i\{\neg\phi : \phi \in O_i\} \cup \bigcup_i\{(\phi\wedge\psi) : \phi,\psi \in O_i\} && \text{(A.27)}\\
&= \bigcup\{\Theta \cup \{\triangle_i T : T \in T_i\} \cup \{\triangledown_i T : T \in T_i\}\cup \\
&\qquad \{\neg\phi : \phi \in O_i\} \cup \{(\phi\wedge\psi) : \phi,\psi \in O_i\} : \langle T_i, O_i, E_i\rangle \in C && \text{(A.28)}\\
&= \bigcup\{g_O(c) : c \in C\} && \text{(A.29)}
\end{aligned}
$$

Third, we show that

$$
g_E(\bigsqcup_{\sqsubset} C) = \bigcup\{g_E(c) : c \in C\} \tag{A.30}
$$

$$
\begin{aligned}
g_E(\bigsqcup_{\sqsubset} C) &= g_E(\bigcup_i\{T_i\},\bigcup_i\{O_i\},\bigcup_i\{E_i\}) && \text{By (A.12)}\\
&=\bigcup_i\{O_i\} \cup \{T \doteq U : T,U \in \bigcup_i\{T_i\}\}\cup \\
&\quad \{\neg\phi : \phi \in \bigcup_i\{E_i\}\} \cup \{(\phi\wedge\psi) : \phi,\psi \in \bigcup_i\{E_i\}\} && \text{(A.31)}\\
&=\bigcup_i\{O_i\} \cup \{T \doteq U : T,U \in T_j \text{for some } T_j\}\cup \\
&\quad \{\neg\phi : \phi \in \bigcup_i\{E_i\}\} \cup \{(\phi\wedge\psi) : \phi,\psi \in E_j \text{for some } E_j\} && \text{(A.32)}\\
&=\bigcup_i\{O_i\} \cup \bigcup_i\{T \doteq U : T,U \in T_i\}\cup \\
&\quad \bigcup_i\{\neg\phi : \phi \in E_i\} \cup \bigcup_i\{(\phi\wedge\psi) : \phi,\psi \in E_i\} && \text{(A.33)}\\
&= \bigcup\{O_i \cup \{T \doteq U : T,U \in T_i\}\cup \\
&\qquad \{\neg\phi : \phi \in E_i\} \cup \{(\phi\wedge\psi) : \phi,\psi \in E_i\} : \langle T_i, O_i, E_i\rangle \in C\} && \text{(A.34)}\\
&= \bigcup\{g_E(c) : c \in C\} && \text{(A.35)}
\end{aligned}
$$

Finally, we show eq. A.13:

$$
\begin{aligned}
g(\bigsqcup_{\sqsubset} C) &= \langle g_T(\bigsqcup_{\sqsubset} C), g_O(\bigsqcup_{\sqsubset} C), g_E(\bigsqcup_{\sqsubset} C)\rangle && \text{(A.36)}\\
&= \langle \bigcup\{g_T(c) : c \in C\}, \bigcup\{g_O(c) : c \in C\}, \bigcup\{g_E(c) : c \in C\}\rangle && \text{(A.37)}\\
&\qquad \text{By (A.16),(A.25),(A.30)} && \text{(A.38)}\\
&= \bigsqcup_{\sqsubset}\{g(c) : c \in C\} && \text{By (A.12)}
\end{aligned}
$$

Thus g has a (least) fixed point, which is a solution to the recursive definition.

■

Corollary A.3 (Fixed Point Characterization of *TL*, *AL* and *EL*)

$$\langle TL, AL, EL\rangle = \bigsqcup_{\sqsubset}\{g^n(\emptyset, \emptyset, \emptyset) : n \geq 0\} \tag{A.39}$$

■

Appendix B

Proof of Lemma 4.2

To save space, we only show proofs for selected parts of the lemma. Formally, Corollary 4.9 shows that the term calculus is complete and hence that the rest of the lemma also holds. All parts of the lemma which is used in the proof of Corollary 4.9 are among the selected parts proven below.

1.

1	$\vdash T \sqcap (T \sqcup U) \doteq T$	T11
2	$\vdash T \doteq T \sqcap (T \sqcup U)$	**MP**(T2,1)
3	$\vdash T \doteq T$	T1
4	$\vdash T \sqcup T \doteq T \sqcup (T \sqcap (T \sqcup U))$	**MP**(T4,2,3)
5	$\vdash T \sqcup (T \sqcap (T \sqcup U)) \doteq T$	T10
6	$\vdash T \sqcup T \doteq T \sqcup (T \sqcap (T \sqcup U)) \wedge T \sqcup (T \sqcap (T \sqcup U)) \doteq T$	**MP**(**Prop**,4,5)
7	$\vdash T \sqcup T \doteq T$	**MP**(T3,6)

2.

1	$\vdash T \sqcup (T \sqcap U) \doteq T$	T10
2	$\vdash T \doteq T \sqcup (T \sqcap U)$	**MP**(T2,1)
3	$\vdash T \doteq T$	T1
4	$\vdash T \sqcap T \doteq T \sqcap (T \sqcup (T \sqcap U))$	**MP**(T5,2,3)
5	$\vdash T \sqcap (T \sqcup (T \sqcap U)) \doteq T$	T11
6	$\vdash T \sqcap T \doteq T \sqcap (T \sqcup (T \sqcap U)) \wedge T \sqcap (T \sqcup (T \sqcap U)) \doteq T$	**MP**(**Prop**,4,5)
7	$\vdash T \sqcap T \doteq T$	**MP**(T3,6)

12.

1	$\vdash T \doteq U \sqcup T \wedge U \sqcup T \doteq U \to T \doteq U$	T3
2	$\vdash U \sqcup T \doteq T \to T \doteq U \sqcup T$	T2
3	$\vdash T \sqcup U \doteq U \to U \sqcup T \doteq U$	**MP**(**Prop**,T4,T6)
4	$\vdash T \sqcup U \doteq U \wedge U \sqcup T \doteq T \to U \sqcup T \doteq U \wedge T \doteq U \sqcup T$	**MP**(**Prop**,2,3)
5	$\vdash T \sqcup U \doteq U \wedge U \sqcup T \doteq T \to T \doteq U$	**MP**(T3,1,4)
6	$\vdash T \doteq U \to T \sqcup U \doteq U \sqcup U$	**MP**(**Prop**,T4,T1)
7	$\vdash U \sqcup U \doteq U$	L.4.2.1
8	$\vdash T \doteq U \to T \sqcup U \doteq U$	**MP**(**Prop**,T3,6,7)
9	$\vdash U \doteq T \to U \sqcup T \doteq T \sqcup T$	**MP**(**Prop**,T4,T1)
10	$\vdash T \sqcup T \doteq T$	L.4.2.1
11	$\vdash U \doteq T \to U \sqcup T \doteq T$	**MP**(**Prop**,T3,9,10)
12	$\vdash T \doteq U \to U \doteq T$	T2
13	$\vdash T \doteq U \to U \sqcup T \doteq T$	**MP**(T3,11,12)
14	$\vdash T \doteq U \leftrightarrow T \sqcup U \doteq U \wedge U \sqcup T \doteq T$	**MP**(**Prop**,5,8,13)

13.

1	$\vdash (T \sqcup U) \sqcap (T \sqcup V) \doteq ((T \sqcup U) \sqcap T) \sqcup ((T \sqcup U) \sqcap V)$	T12
2	$\vdash T \sqcap (T \sqcup U) \doteq T$	T11
3	$\vdash (T \sqcup U) \sqcap V \doteq V \sqcap (T \sqcup U)$	T7
4	$\vdash V \sqcap (T \sqcup U) \doteq (V \sqcap T) \sqcup (V \sqcap U)$	T12
5	$\vdash (T \sqcup U) \sqcap V \doteq (V \sqcap T) \sqcup (V \sqcap U)$	**MP**(T3,3,4)
6	$\vdash ((T \sqcup U) \sqcap T) \sqcup ((T \sqcup U) \sqcap V) \doteq T \sqcup ((V \sqcap T) \sqcup (V \sqcap U))$	**MP**(T4,2,5)
7	$\vdash (T \sqcup U) \sqcap (T \sqcup V) \doteq T \sqcup ((V \sqcap T) \sqcup (V \sqcap U))$	**MP**(T3,1,6)
8	$\vdash T \sqcup ((V \sqcap T) \sqcup (V \sqcap U)) \doteq (T \sqcup (V \sqcap T)) \sqcup (V \sqcap U)$	T8
9	$\vdash (T \sqcup U) \sqcap (T \sqcup V) \doteq (T \sqcup (V \sqcap T)) \sqcup (V \sqcap U)$	**MP**(T3,7,8)
10	$\vdash V \sqcap T \doteq T \sqcap V$	T2
11	$\vdash T \doteq T$	T1
12	$\vdash T \sqcup (V \sqcap T) \doteq T \sqcup (T \sqcap V)$	**MP**(T4,10,11)
13	$\vdash T \sqcup (T \sqcap V) \doteq T$	T10
14	$\vdash T \sqcup (V \sqcap T) \doteq T$	**MP**(T3,12,13)
15	$\vdash V \sqcap U \doteq U \sqcap V$	T2
16	$\vdash (T \sqcup (V \sqcap T)) \sqcup (V \sqcap U) \doteq T \sqcup (U \sqcap V)$	**MP**(T4,14,15)
17	$\vdash (T \sqcup U) \sqcap (T \sqcup V) \doteq T \sqcup (U \sqcap V)$	**MP**(T3,9,16)
18	$\vdash T \sqcup (U \sqcap V) \doteq (T \sqcup U) \sqcap (T \sqcup V)$	**MP**(T2,17)

Appendix C

Proof of Lemma 4.7

We prove both parts of the lemma simultaneously, by induction over the degree of $T_1 \sqcup T_2$ (of course, $d(T_1) \leq d(T_1 \sqcup T_2) \geq d(T_2)$).

Induction basis (Lemma 4.7) For the induction basis, let $d(T_1 \sqcup T_2) = 1$ – in which case $d(T_1) = d(T_2) = 1$. We first show an intermediate result:

$$\text{If } d(\{\alpha\}) = d(\{\beta\}) = 1 \text{ and } [\alpha] \neq [\beta] \text{ then } \vdash \neg(\{\alpha\} \doteq \{\beta\}) \tag{C.1}$$

Let $d(\{\alpha\}) = d(\{\beta\}) = 1$ and $[\alpha] \neq [\beta]$. We show (C.1) by structural induction over α. Note that since $\{\alpha\}$ and $\{\beta\}$ are of degree 1, neither α or β can contain subformulas $\triangle_i T$ or $\bigtriangledown_i T$.

Induction base (C.1) For the single base case, let $\alpha = p$. Then $[\alpha] = p$. Assume that $[\beta] \neq p$. Then $\beta \neq p$, and

$$\vdash \neg(\{\alpha\} \doteq \{\beta\}) \tag{C.2}$$

holds by N2.

Induction step (C.1) For the induction step, first let $\alpha = \neg\gamma$ – then $d(\{\gamma\}) = 1$ and $[\alpha] = \neg[\gamma]$. The induction hypothesis is that, for all γ' with $d(\{\gamma'\}) = 1$, if $[\gamma] \neq [\gamma']$ then $\vdash \neg(\{\gamma\} \doteq \{\gamma'\})$. $[\beta] \neq \neg[\gamma]$. If β starts with a negation, $\beta = \neg\beta'$ where $d(\{\beta'\}) = 1$ and $[\gamma] \neq [\beta']$. Then $\vdash \neg(\{\gamma\} \doteq \{\beta'\})$ by the induction hypothesis, and (C.2) follows by N7. If β does not start with a negation, (C.2) follows by N8. For the second part of the induction step, let $\alpha = (\gamma_1 \wedge \gamma_2)$. Then $[\alpha] = ([\gamma_1] \wedge [\gamma_2])$. $[\beta] \neq ([\gamma_1] \wedge [\gamma_2])$. If β is not a conjunction, then (C.2) follows by N10. If $\beta = (\beta_1 \wedge \beta_2)$ then $([\gamma_1] \wedge [\gamma_2]) \neq ([\beta_1] \wedge [\beta_2])$. Then either $[\gamma_1] \neq [\beta_1]$ or $[\gamma_2] \neq [\beta_2]$ and either $\vdash \neg(\{\gamma_1\} \doteq \{\beta_1\})$ or $\vdash \neg(\{\gamma_2\} \doteq \{\beta_2\})$ holds by the induction hypothesis, and then (C.2) follows by N9.

We now show that, if $\{\beta_j\}$ $(1 \leq j \leq l)$ are terms,

$$\text{If } d(T) = 1 \text{ and } [T] = \{[\beta_1], \ldots, [\beta_l]\} \text{ then } \vdash T \doteq \{\beta_1, \ldots, \beta_l\} \tag{C.3}$$

by structural induction over T. We make use of the fact that for any term S of degree 1, if $[S] = \{[\alpha_1], \ldots, [\alpha_k]\}$ then $[\alpha_j] = \alpha_j$ $(1 \leq j \leq k)$ (because $d(\{\alpha_j\}) = 1$; α_j cannot contain any knowledge operators) and thus $[S] = \{\alpha_1, \ldots, \alpha_k\}$[1]. Let $[\mathrm{T}] = \{[\beta_1], \ldots, [\beta_l]\} = \{\beta_1, \ldots, \beta_l\}$.

[1]Recall that S is a *term* and an expression like $S = \{\alpha_1, \ldots, \alpha_k\}$ is a shorthand for $S = \underline{\{}\alpha_1\underline{,} \ldots \underline{,}\alpha_k\underline{\}}$, while $[S]$ is a *set* and an expression like $[S] = \{\alpha_1, \ldots, \alpha_k\}$ denotes the set consisting of the elements $\alpha_1, \ldots, \alpha_k$. It is only a formula which does not contain a knowledge operator, e.g. all formulae in a term of degree 1, that has an interpretation equal to the formula itself.

Induction basis (C.3) For the induction basis, let $T = \{\alpha_1, \ldots, \alpha_k\}$. Then $\{[\beta_1], \ldots, [\beta_l]\} = [T] = \{[\alpha_1], \ldots, [\alpha_k]\}$, $\{\beta_1, \ldots, \beta_l\} = \{\alpha_1, \ldots, \alpha_k\}$ and $\vdash \{\alpha_1, \ldots, \alpha_k\} \doteq \{\beta_1, \ldots, \beta_l\}$ by Lemma 4.3.

Induction step (C.3) For the induction step, first consider the case when $T = U \sqcup V$. $[U] \cup [V] = \{[\beta_1], \ldots, [\beta_l]\}$, so $[U] = \{[\beta_1^U], \ldots, [\beta_{k_U}^U]\}$ and $[V] = \{[\beta_1^V], \ldots, [\beta_{k_V}^V]\}$, where each $\beta_j^U, \beta_j^V \in \{\beta_1, \ldots, \beta_l\}$.

1	$\vdash U \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U\}$	Ind. hyp.
2	$\vdash V \doteq \{\beta_1^V, \ldots, \beta_{k_V}^V\}$	Ind. hyp.
3	$\vdash U \sqcup V \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup V$	**Repl**,1
4	$\vdash \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup V \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup \{\beta_1^V, \ldots, \beta_{k_V}^V\}$	**Repl**,2
5	$\vdash U \sqcup V \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup \{\beta_1^V, \ldots, \beta_{k_V}^V\}$	**MP**,T3,3,4

From several applications of T13, and **Repl**, we have that

6	$\vdash \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup \{\beta_1^V, \ldots, \beta_{k_V}^V\} \doteq$ $\{\beta_1^U, \ldots, \beta_{k_U}^U, \beta_1^V, \ldots, \beta_{k_V}^V\}$	
7	$\vdash U \sqcup V \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U, \beta_1^V, \ldots, \beta_{k_V}^V\}$	**MP**,T3,5,6

The sequence $\beta_1^U, \ldots, \beta_{k_U}^U, \beta_1^V, \ldots, \beta_{k_V}^V$ is a permutation, possibly with duplicates, of the sequence $\beta_1, \ldots, \beta_l$, so

$$\vdash U \sqcup V \doteq \{\beta_1, \ldots, \beta_l\}$$

by Lemma 4.3 and T3. For the second case in the induction step, let $T = U \sqcap V$ where $[U] = \{[\alpha_1^U], \ldots, [\alpha_{k_U}^U]\} = \{\alpha_1^U, \ldots, \alpha_{k_U}^U\}$ and $[V] = \{[\alpha_1^V], \ldots, [\alpha_{k_V}^V]\} = \{\alpha_1^V, \ldots, \alpha_{k_V}^V\}$. In order to prove this case, we first show that

$$\vdash \{\alpha_1^U, \ldots, \alpha_{k_U}^U\} \sqcap \{\alpha_1^V, \ldots, \alpha_{k_V}^V\} \doteq \{\beta_1, \ldots, \beta_l\} \tag{C.4}$$

We have that

$$\begin{aligned}
\vdash \{\alpha_1^U, \ldots, \alpha_{k_U}^U\} \sqcap \{\alpha_1^V, \ldots, \alpha_{k_V}^V\} \doteq \\
(\{\alpha_1^U\} \sqcap \{\alpha_1^V\}) \sqcup (\{\alpha_1^U\} \sqcap \{\alpha_2^V\}) \sqcup \cdots \sqcup (\{\alpha_1^U\} \sqcap \{\alpha_{k_V}^V\}) \sqcup \\
(\{\alpha_2^U\} \sqcap \{\alpha_1^V\}) \sqcup (\{\alpha_2^U\} \sqcap \{\alpha_2^V\}) \sqcup \cdots \sqcup (\{\alpha_2^U\} \sqcap \{\alpha_{k_V}^V\}) \sqcup \\
\vdots \\
(\{\alpha_{k_U}^U\} \sqcap \{\alpha_1^V\}) \sqcup (\{\alpha_{k_U}^U\} \sqcap \{\alpha_2^V\}) \sqcup \cdots \sqcup (\{\alpha_{k_U}^U\} \sqcap \{\alpha_{k_V}^V\})
\end{aligned} \tag{C.5}$$

by T13 and repeated applications of T12, **Repl** and the axioms of equality. The above expression contains exactly one term on the form $\{\alpha^U\} \sqcap \{\alpha^V\}$ for each pair $\alpha^U \in [U], \alpha^V \in [V]$. If $\beta \in [T]$, then $\alpha_i^U = \beta \in [U]$ and $\alpha_j^V = \beta \in [V]$ and we have $\vdash \{\alpha_i^U\} \doteq \{\alpha_j^V\}$ by T1 and thus $\vdash \{\alpha_i^U\} \sqcap \{\alpha_j^V\} \doteq \{\beta\}$ by T14. If $\alpha_i^U \in [U]$ but $\alpha_i^U \notin [T]$, for every $\alpha_j^V \in [V]$ $\alpha_i^U \neq \alpha_j^V$, and because $[\alpha_i^U] = \alpha_i^U$ and $[\alpha_j^V] = \alpha_j^V$ ($\{\alpha_i^U\}, \{\alpha_j^V\}$ are of degree 1), $\vdash \neg(\{\alpha_i^U\} \doteq \{\alpha_j^V\})$ by (C.1) and thus $\vdash \{\alpha_i^U\} \sqcap \{\alpha_j^V\} \doteq \emptyset$ by T15. Similarly, if $\alpha_j^V \in [V]$ but $\alpha_j^V \notin [T]$, $\vdash \{\alpha_i^U\} \sqcap \{\alpha_j^V\} \doteq \emptyset$ for every $\alpha_i^U \in [U]$. Then, (C.4) follows by (C.5), **Repl** and T13. We now use (C.4) to show the current induction step:

1	$\vdash U \doteq \{\alpha_1^U, \dots, \alpha_{k_U}^U\}$	Ind. hyp.
2	$\vdash V \doteq \{\alpha_1^V, \dots, \alpha_{k_V}^V\}$	Ind. hyp.
3	$\vdash U \sqcap V \doteq \{\alpha_1^U, \dots, \alpha_{k_U}^U\} \sqcap \{\alpha_1^V, \dots, \alpha_{k_V}^V\}$	**MP**,T5,1,2
4	$\vdash \{\alpha_1^U, \dots, \alpha_{k_U}^U\} \sqcap \{\alpha_1^V, \dots, \alpha_{k_V}^V\} \doteq \{\beta_1, \dots, \beta_l\}$	(C.4)
5	$\vdash U \sqcap V \doteq \{\beta_1, \dots, \beta_l\}$	**MP**,T3,3,4

This completes the proof of (C.3).

We can now show the base case in the inductive proof over the degree of $T_1 \sqcup T_2$. Let α_j, β_j be such that $[T_1] = \{[\alpha_1], \dots, [\alpha_k]\}$ and $[T_2] = \{[\beta_1], \dots, [\beta_m]\}$. First, consider the case that $[T_1] = [T_2]$:

1	$\vdash T_1 \doteq \{\alpha_1, \dots, \alpha_k\}$	(C.3)
2	$\vdash T_2 \doteq \{\alpha_1, \dots, \alpha_k\}$	(C.3)
3	$\vdash \{\alpha_1, \dots, \alpha_k\} \doteq T_2$	**MP**,T2,2
4	$\vdash T_1 \doteq T_2$	**MP**,T3,1,3

Second, consider the case that $[T_1] \neq [T_2]$. We show that

$$\vdash \neg(\{\alpha_1, \dots, \alpha_k\} \doteq \{\beta_1, \dots, \beta_m\}) \tag{C.6}$$

We assume that there is an $[\alpha_i] \in [T_1]$ such that $[\alpha_i] \notin [T_2]$ (the proof is equivalent in the case that $[\beta_i] \in [T_2]$ and $[\beta_i] \notin [T_1]$). Since $[\alpha_i] \neq [\beta_j]$ $(1 \leq j \leq m)$, $\vdash \neg(\{\alpha_i\} \doteq \{\beta_j\})$ by (C.1) and $\vdash \neg((\{\alpha_i\} \doteq \{\beta_1\}) \vee \cdots \vee (\{\alpha_i\} \doteq \{\beta_m\}))$ by **Prop**. By **Prop** and N1, $\vdash \neg(\{\alpha_1, \dots, \alpha_k\} \preceq \{\beta_1, \dots, \beta_m\})$ and (C.6) follows by Lemma 4.2.12. By (C.3) $\vdash \{\alpha_1, \dots, \alpha_k\} \doteq T_1$, and by (C.6), T3 and **Prop**

$$\vdash \neg(T_1 \doteq \{\beta_1, \dots, \beta_m\}) \tag{C.7}$$

By (C.3) $\vdash \{\beta_1, \dots, \beta_m\} \doteq T_2$, and by (C.7), T3, **Prop** and T2,

$$\vdash \neg T_1 \doteq T_2$$

This completes the base case in the inductive proof of the lemma.

Induction step (Lemma 4.7) For the induction step, let the lemma hold for all terms T_1, T_2 such that $d(T_1 \sqcup T_2) \leq k$. Let $d(T_1 \sqcup T_2) = k + 1$. We now show two intermediate results. We first show that

$$\text{If } d(\{\alpha\}) \leq k + 1 \geq d(\{\beta\}) \text{ and } [\alpha] = [\beta] \text{ then } \vdash \{\alpha\} \doteq \{\beta\} \tag{C.8}$$

for all β by structural induction over α. Let $d(\{\alpha\}) \leq k + 1 \geq d(\{\beta\})$ and $[\alpha] = [\beta]$.

Induction base (C.8) For the first base case, let $\alpha = p \in \Theta$. Then $\beta = p$, and

$$\vdash \{\alpha\} \doteq \{\beta\} \tag{C.9}$$

follows by T1. For the second base case, let $\alpha = \triangle_i S_1$. Then $\beta = \triangle_i S_2$ where $[S_1] = [S_2]$. $d(S_1) \leq k \geq d(S_2)$, so $d(S_1 \sqcup S_2) \leq k$ and $\vdash S_1 \doteq S_2$ holds by the induction hypothesis (in the "outmost" inductive proof). (C.9) follows by N3. Similarly, in the third base case when $\alpha = \triangledown_i S_1$ and $\beta = \triangledown_i S_2$, (C.9) follows by the induction hypothesis and N5.

Induction step (C.8) For the induction step, first let $\alpha = \neg\gamma$ – then $[\alpha] = \neg[\gamma]$. The induction hypothesis is that, for all γ', if $[\gamma] = [\gamma']$ then $\vdash \{\gamma\} \doteq \{\gamma'\}$. β must start with a negation; say $\beta = \neg\beta'$ where $[\beta'] = [\gamma]$. $\vdash \{\gamma\} \doteq \{\beta'\}$ by the induction hypothesis, and (C.9) follows by N7. For the second part of the induction step, let $\alpha = (\gamma_1 \wedge \gamma_2)$. Then $[\alpha] = ([\gamma_1] \wedge [\gamma_2])$. β must be a conjunction; say $\beta = (\beta_1 \wedge \beta_2)$ where $[\beta_1] = [\gamma_1]$ and $[\beta_2] = [\gamma_2]$. $\vdash \{\gamma_1\} \doteq \{\beta_1\}$ and $\vdash \{\gamma_2\} \doteq \{\beta_2\}$ by the induction hypothesis, and (C.9) follows by N9.

Second, we show a general version of (C.1):

$$\text{If } d(\{\alpha\}) \leq k+1 \geq d(\{\beta\}) \text{ and } [\alpha] \neq [\beta] \text{ then } \vdash \neg\{\alpha\} \doteq \{\beta\} \qquad \text{(C.10)}$$

for all β by structural induction over α. Let $d(\{\alpha\}) \leq k+1 \geq d(\{\beta\})$ and $[\alpha] \neq [\beta]$.

Induction base (C.10) For the first base case, let $\alpha = p$. Then $[\alpha] = p$. Then $\beta \neq p$, and

$$\vdash \neg(\{\alpha\} \doteq \{\beta\}) \qquad \text{(C.11)}$$

follows by N2. For the second base case, let $\alpha = \triangle_i S_1$. Then $[\alpha] = \triangle_i[S_1]$ and $[\beta] \neq \triangle_i[S_1]$. If $\beta \neq \triangle_i S_2$ (for any S_2), then (C.11) follows by N4. If $\beta = \triangle_i S_2$ (for some S_2), then $[S_2] \neq [S_1]$. $d(S_1) \leq k \geq d(S_2)$, so $d(S_1 \sqcup S_2) \leq k$ and $\vdash \neg S_1 \doteq S_2$ holds by the induction hypothesis (in the "outmost" inductive proof). (C.11) follows by N3. Similarly, in the third base case when $\alpha = \triangledown_i S_1$ and $\beta \neq \triangle_i S_2$ (for any S_2) then (C.11) follows by N6, and if $\alpha = \triangledown_i S_1$ and $\beta = \triangledown_i S_2$, (C.11) follows by the induction hypothesis and N5.

Induction step (C.10) For the induction step, first let $\alpha = \neg\gamma$ – then $[\alpha] = \neg[\gamma]$. The induction hypothesis is that, for all γ', if $[\gamma] \neq [\gamma']$ then $\vdash \neg(\{\gamma\} \doteq \{\gamma'\})$. $[\beta] \neq \neg[\gamma]$. If β starts with a negation, $\beta = \neg\beta'$ where $[\gamma] \neq [\beta']$, then $\vdash \neg(\{\gamma\} \doteq \{\beta'\})$ by the induction hypothesis, and $\vdash \neg(\{\neg\gamma\} \doteq \{\neg\beta'\})$ by N7. If β does not start with a negation, (C.11) follows by N8. For the second part of the induction step, let $\alpha = (\gamma_1 \wedge \gamma_2)$. Then $[\alpha] = ([\gamma_1] \wedge [\gamma_2])$. $[\beta] \neq ([\gamma_1] \wedge [\gamma_2])$. If β is not a conjunction, then (C.11) follows by N10. If $\beta = (\beta_1 \wedge \beta_2)$ then $([\gamma_1] \wedge [\gamma_2]) \neq ([\beta_1] \wedge [\beta_2])$. Then either $[\gamma_1] \neq [\beta_1]$ or $[\gamma_2] \neq [\beta_2]$ and either $\vdash \neg(\{\gamma_1\} \doteq \{\beta_1\})$ or $\vdash \neg(\{\gamma_2\} \doteq \{\beta_2\})$ holds by the induction hypothesis, and then (C.11) follows by N9.

As in the base case, we use structural induction over T to prove

$$\text{If } d(T) \leq k+1 \text{ and } [T] = \{[\beta_1], \ldots, [\beta_l]\} \text{ then } \vdash T \doteq \{\beta_1, \ldots, \beta_l\} \qquad \text{(C.12)}$$

Let $d(T) \leq k+1$ and $[T] = \{[\beta_1], \ldots, [\beta_l]\}$. Clearly, $d(\{\beta_j\}) \leq k+1$ $(1 \leq j \leq l)$.

Induction base (C.12) For the base case, let $T = \{\alpha_1, \ldots, \alpha_l\}$ (in the base case T is a basic term, and it must consist of l formulae since there are l elements in $[T]$). $\{[\alpha_1], \ldots, [\alpha_l]\} = \{[\beta_1, \ldots, [\beta_l]\}$, and we assume that $[\alpha_j] = [\beta_j]$ $(1 \leq j \leq l)$ for simplicity (otherwise just change the indices). By T13 $\vdash T \doteq \{\alpha_1\} \sqcup \cdots \sqcup \{\alpha_l\}$, by (C.8) $\vdash \{\alpha_j\} \doteq \{\beta_j\}$, and by repeated applications of **Repl** and T3, $\vdash T \doteq \{\beta_1\} \sqcup \cdots \sqcup \{\beta_l\}$ and thus $\vdash T \doteq \{\beta_1, \ldots, \beta_l\}$ by T13 – which is what we needed to show for the base in the structural induction over T.

Induction step (C.12) For the induction step, consider the case when $T = U \sqcup V$. In this case the proof of (C.12) is identical to the corresponding proof in the base case ($d(T \sqcup U) = 1$): Let $[U] = \{[\beta_1^U], \ldots, [\beta_{k_U}^U]\}$ and $[V] = \{[\beta_1^V], \ldots, [\beta_{k_V}^V]\}$, where each $\beta_j^U, \beta_j^V \in \{\beta_1, \ldots, \beta_l\}$.

1	$\vdash U \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U\}$	Ind. hyp.
2	$\vdash V \doteq \{\beta_1^V, \ldots, \beta_{k_V}^V\}$	Ind. hyp.
3	$\vdash U \sqcup V \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup V$	**Repl**,1
4	$\vdash \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup V \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup \{\beta_1^V, \ldots, \beta_{k_V}^V\}$	**Repl**,2
5	$\vdash U \sqcup V \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup \{\beta_1^V, \ldots, \beta_{k_V}^V\}$	**MP**,T3,3,4

From several applications of T13, and **Repl**, we have that

6	$\vdash \{\beta_1^U, \ldots, \beta_{k_U}^U\} \sqcup \{\beta_1^V, \ldots, \beta_{k_V}^V\} \doteq$ $\{\beta_1^U, \ldots, \beta_{k_U}^U, \beta_1^V, \ldots, \beta_{k_V}^V\}$	
7	$\vdash U \sqcup V \doteq \{\beta_1^U, \ldots, \beta_{k_U}^U, \beta_1^V, \ldots, \beta_{k_V}^V\}$	**MP**,T3,5,6

The sequence $\beta_1^U, \ldots, \beta_{k_U}^U, \beta_1^V, \ldots, \beta_{k_V}^V$ is a permutation, possibly with duplicates, of the sequence $\beta_1, \ldots, \beta_l$, so

$$\vdash U \sqcup V \doteq \{\beta_1, \ldots, \beta_l\}$$

by Lemma 4.3 and T3. For the second case in the induction step, let $T = U \sqcap V$ where $[U] = \{[\alpha_1^U], \ldots, [\alpha_{k_U}^U]\}$ and $[V] = \{[\alpha_1^V], \ldots, [\alpha_{k_V}^V]\}$. Since $d(U) \leq k+1 \geq d(V)$, $d(\{\alpha_{j_U}^U\}) \leq k+1 \geq d(\{\alpha_{j_V}^V\})$ ($1 \leq j_U \leq k_U$, $1 \leq j_V \leq k_V$). We proceed exactly as in the base case ($d(T \sqcup U) = 1$) by first proving (C.4):

$$\vdash \{\alpha_1^U, \ldots, \alpha_{k_U}^U\} \sqcap \{\alpha_1^V, \ldots, \alpha_{k_V}^V\} \doteq \{\beta_1, \ldots, \beta_l\} \tag{C.13}$$

in the same way as in the base case. Again, we have that

$$\begin{aligned}\vdash \{\alpha_1^U, \ldots, \alpha_{k_U}^U\} \sqcap \{\alpha_1^V, \ldots, \alpha_{k_V}^V\} \doteq \\ (\{\alpha_1^U\} \sqcap \{\alpha_1^V\}) \sqcup (\{\alpha_1^U\} \sqcap \{\alpha_2^V\}) \sqcup \cdots \sqcup (\{\alpha_1^U\} \sqcap \{\alpha_{k_V}^V\}) \sqcup \\ (\{\alpha_2^U\} \sqcap \{\alpha_1^V\}) \sqcup (\{\alpha_2^U\} \sqcap \{\alpha_2^V\}) \sqcup \cdots \sqcup (\{\alpha_2^U\} \sqcap \{\alpha_{k_V}^V\}) \sqcup \\ \vdots \\ (\{\alpha_{k_U}^U\} \sqcap \{\alpha_1^V\}) \sqcup (\{\alpha_{k_U}^U\} \sqcap \{\alpha_2^V\}) \sqcup \cdots \sqcup (\{\alpha_{k_U}^U\} \sqcap \{\alpha_{k_V}^V\})\end{aligned} \tag{C.14}$$

by T13 and repeated applications of T12, **Repl** and the axioms of equality. The above expression contains exactly one term on the form $\{\alpha^U\} \sqcap \{\alpha^V\}$ for each pair $[\alpha^U] \in [U], [\alpha^V] \in [V]$. If $[\beta] \in [T]$, then $[\alpha_h^U] = [\beta] \in [U]$ and $[\alpha_j^V] = [\beta] \in [V]$, for some h, j. Since $d(\{\alpha_h^V\}), d(\{\alpha_j^V\}), d(\{\beta\}) \leq k+1$, $\vdash \{\alpha_j^V\} \doteq \{\beta\}$ and $\vdash \{\alpha_h^U\} \doteq \{\beta\}$ by (C.8) and $\vdash \{\alpha_h^U\} \sqcap \{\alpha_j^V\} \doteq \{\alpha_h^U\} \sqcap \{\beta\}$ by **Repl**. By T14 $\vdash \{\beta\} \sqcap \{\alpha_h^U\} \doteq \{\beta\}$, and thus $\vdash \{\alpha_h^U\} \sqcap \{\alpha_j^V\} \doteq \{\beta\}$, by equational calculus. If $[\alpha_h^U] \in [U]$ but $[\alpha_h^U] \notin [T]$, then, for every $[\alpha_j^V] \in [V]$, $[\alpha_h^U] \neq [\alpha_j^V]$, and, since $d(\{\alpha_h^U\}) \leq k+1 \geq d(\{\alpha_j^V\})$, $\vdash \neg(\{\alpha_h^U\} \doteq \{\alpha_j^V\})$ by (C.10) and thus $\vdash \{\alpha_h^U\} \sqcap \{\alpha_j^V\} \doteq \emptyset$ by T15. Similarly, if $[\alpha_j^V] \in [V]$ but $[\alpha_j^V] \notin [T]$, $\vdash \{\alpha_h^U\} \sqcap \{\alpha_j^V\} \doteq \emptyset$ for every $[\alpha_h^U] \in [U]$. Thus, (C.13) follows by (C.14),

Repl and T13. We now use (C.13) to show the current induction step:

1	$\vdash U \doteq \{\alpha_1^U, \ldots, \alpha_{k_U}^U\}$	Ind. hyp.
2	$\vdash V \doteq \{\alpha_1^V, \ldots, \alpha_{k_V}^V\}$	Ind. hyp.
3	$\vdash U \sqcap V \doteq \{\alpha_1^U, \ldots, \alpha_{k_U}^U\} \sqcap \{\alpha_1^V, \ldots, \alpha_{k_V}^V\}$	**MP**,T5,1,2
4	$\vdash \{\alpha_1^U, \ldots, \alpha_{k_U}^U\} \sqcap \{\alpha_1^V, \ldots, \alpha_{k_V}^V\} \doteq \{\beta_1, \ldots, \beta_l\}$	(C.13)
5	$\vdash U \sqcap V \doteq \{\beta_1, \ldots, \beta_l\}$	**MP**,T3,3,4

This completes the proof of (C.12).

We can now use (C.12) show the the inductive step in the proof of the lemma, in exactly the same way we used (C.3) and (C.6) in the base case: Clearly, there exists α_j, β_j such that $[T_1] = \{[\alpha_1], \ldots, [\alpha_k]\}$ and $[T_2] = \{[\beta_1], \ldots, [\beta_m]\}$. First, consider the case that $[T_1] = [T_2]$:

1	$\vdash T_1 \doteq \{\alpha_1, \ldots, \alpha_k\}$	(C.12)
2	$\vdash T_2 \doteq \{\alpha_1, \ldots, \alpha_k\}$	(C.12)
3	$\vdash \{\alpha_1, \ldots, \alpha_k\} \doteq T_2$	**MP**,T2,2
4	$\vdash T_1 \doteq T_2$	**MP**,T3,1,3

Second, consider the case that $[T_1] \neq [T_2]$. As in the base case, we show that:

$$\vdash \neg(\{\alpha_1, \ldots, \alpha_k\} \doteq \{\beta_1, \ldots, \beta_m\}) \tag{C.15}$$

We assume that there is an $[\alpha_i] \in [T_1]$ such that $[\alpha_i] \notin [T_2]$ (the proof is equivalent in the case that $[\beta_i] \in [T_2]$ and $[\beta_i] \notin [T_1]$). Since $[\alpha_i] \neq [\beta_j]$ and $d(\{\alpha_i\}) \leq k+1 \geq d(\{\beta_j\})$ $(1 \leq j \leq m)$, $\vdash \neg(\{\alpha_i\} \doteq \{\beta_j\})$ by (C.10) and $\vdash \neg((\{\alpha_i\} \doteq \{\beta_1\}) \vee \cdots \vee (\{\alpha_i\} \doteq \{\beta_m\}))$ by **Prop**. By **Prop** and N1, $\vdash \neg(\{\alpha_1, \ldots, \alpha_k\} \preceq \{\beta_1, \ldots, \beta_m\})$ and (C.15) follows by Lemma 4.2.12.

By (C.12) $\vdash \{\alpha_1, \ldots, \alpha_k\} \doteq T_1$, and by (C.15), T3 and **Prop**

$$\vdash \neg(T_1 \doteq \{\beta_1, \ldots, \beta_m\}) \tag{C.16}$$

By (C.12) $\vdash \{\beta_1, \ldots, \beta_m\} \doteq T_2$, and by (C.16), T3, **Prop** and T2,

$$\vdash \neg T_1 \doteq T_2$$