

# Reasoning about Knowledge and Belief: A Syntactical Treatment

Maria Fasli

University of Essex, Department of Computer Science

Wivenhoe Park, Colchester CO4 3SQ, UK

Email: mfasli@essex.ac.uk

## Abstract

The study of formal theories of agents has intensified over the last couple of decades, since such formalisms can be viewed as providing the specifications for building rational agents and multi-agent systems. Most of the proposed approaches are based upon the well-understood framework of modal logics and possible world semantics. Although intuitive and expressive, these approaches lack two properties that can be considered important to a rational agent's reasoning: quantification over the propositional attitudes, and self-referential statements. This paper presents an alternative framework which is different from those found in the literature in two ways: Firstly, a syntactical approach for the representation of the propositional attitudes is adopted. This involves the use of a truth predicate and syntactic modalities which are defined in terms of the truth predicate itself and corresponding modal operators. Secondly, an agent's information state includes both knowledge and beliefs. Independent modal operators for the two notions are introduced and based on them syntactic modalities are defined. Furthermore, the relation between knowledge and belief is thoroughly explored and three different connection axiomatisations for the modalities and the syntactic modalities are proposed and their properties investigated.

**Keywords:** Syntactical Theories, Logics of Knowledge and Belief

## 1 Introduction

The study of knowledge, belief, truth and reasoning dates back to the time of Aristotle in Ancient Greece. More recently, one of the earliest and most well-known treatments of reasoning about knowledge and belief is the philosophical work of Hintikka [25] in the 1960s. Since then, reasoning about the epistemic and doxastic notions has stimulated interest and found applications in such diverse fields as Linguistics, Game Theory and Artificial Intelligence. In Artificial Intelligence and Multi-agent systems in particular, the main objective is to create computational agents that will be able to exhibit "smart" behaviour. However, to achieve such an aim, these complex systems need to be endowed with information about the world they inhabit and the ability to reason with and about this information.

Formalising theories that would describe the behaviour of an intelligent agent is a non-trivial task. A convenient way of determining such a complex system's behaviour is by abstracting away from its mechanistic and design details and considering it to be an intentional system [9]. Hence, an agent is ascribed mental attitudes and characterised as having

a “cognitive state”. Although there is no consensus as to what are the right ingredients of an agent’s mental state, it is generally agreed that it should include attitudes representing the agent’s information state such as knowledge and beliefs, as well as pro-attitudes such as desires, wishes and intentions.

A number of theoretical models for reasoning agents have been proposed [5, 11, 47, 56, 57]. Most of these formalise the information state of an agent in terms of either knowledge or belief. However an adequate theory of agents should include both these propositional attitudes [22]. There are a number of reasons for this. Firstly, the nature of knowledge is different from that of belief, since the former is intimately connected with the notion of truth, whereas the latter does not have this kind of stipulation. Thus, certain things can be accepted as beliefs (e.g. a belief that it will be sunny tomorrow), but cannot be described as knowledge. Secondly, an agent’s decisions which actions to perform in order to achieve its goals need to be based upon information about the world. It seems reasonable to suggest that knowledge carries a different weight in the decision making process of an agent than that of beliefs. Intention formation and its relation to beliefs has been investigated, but its relation and interdependence with knowledge has not been given much attention. Only a few attempts have been made to formalise both information attitudes in the same system and axiomatically capture their relation [31, 54, 55, 33, 42, 58]. Often knowledge is introduced as true belief or justified true belief [5], but no explicit reference is being made to the relationship between knowledge and the other attitudes or its role in the agents’ decision making.

A highly desirable feature for a theory of reasoning agents is the ability to quantify over the objects of the propositional attitudes and express self-referential statements. This stems from the fact that a rational agent in a multi-agent world needs to be able to reason about the world and what holds true as well as about its own cognitive state and that of other agents. This is intimately connected with the ability of an agent to refer to “oneself” in a number of contexts either directly or indirectly and the ability to quantify over propositions. Consider the following statements:

John believes things which are false

Carmen and John share the same desires

Most of the current approaches are based on propositional or first order modal logics with possible worlds semantics [5, 11, 47, 56, 57, 60]. Though expressive, these frameworks lack the ability to express the above statements since quantification over propositions is not allowed. While this is beyond the expressive power of classical first order and modal logics, it can be expressed in higher order logics. For instance, the above statements could be expressed in higher intensional logic [39] as follows:

$$\begin{aligned} &\exists x_{\langle w,t \rangle} (B_{John}(x) \wedge \neg(x)) \\ &\forall x_{\langle w,t \rangle} (D_{Carmen}(x) \Leftrightarrow D_{John}(x)) \end{aligned}$$

But higher order logics are typed and thus the ability to analyse self-referential statements is lost. An alternative would be to use a syntactical theory. Choosing syntactical theories over modal logics has certain advantages. Firstly, they can be implemented in first order logic, and secondly, it seems quite natural to be able to represent the information, motivation and deliberation attitudes of an agent as sentences in some language. The expressive power of the syntactical approach facilitates quantification over propositions in an elegant way. Their main disadvantage is that they are prone to inconsistency.

This paper focuses on two issues. Firstly, a syntactical approach for the representation of the different elements of an agent’s cognitive state is advocated. Secondly, both information-

bearing attitudes of knowledge and belief are formalised. Moreover, connection axiomatisations between knowledge and belief are explored that do not allow the two notions to collapse. The paper is logically divided into two parts: The first part involves the construction of the framework and includes the presentation of the language and semantics. The modal operators and syntactic modalities for knowledge and belief are introduced here. Logics for truth and modalities, and truth and syntactic modalities are also discussed. In the second part the relation between knowledge and belief is considered. Connection axiomatisations that do not allow their collapse are investigated. These are considered in the context of both modalities and syntactic modalities. The paper ends with a discussion of further work, a summary and the conclusions.

## 2 Propositional Attitudes and Propositions

Despite its usefulness and convenience the use of the intentional stance raises, apart from philosophical objections regarding the legitimacy of attributing mental characteristics to machines [36], the important issue of what are the objects of the intentional notions. What is it that is believed, known or intended? More often the answer given to this question is that the objects of the intentional notions are propositions. However, this only creates new questions about the nature of propositions and there seems to be no consensus among philosophers, logicians and researchers in Artificial Intelligence on that. Nevertheless there seem to be two main views on the nature of propositions: the syntactical and the semantical. The first regards propositions as sentence-like entities, whereas the second argues that propositions are not sentences in a language, but their meaning or the represented state of affairs or possible worlds. According to the former, propositions can be constructed from parts according to a syntax. According to the latter, propositions are sets of possible worlds, and two sentences express the same proposition just in case they are true in exactly the same possible worlds. Both views have raised considerable objections.

In the possible worlds conceptualisation propositions are taken to be sets of possible worlds and properties are understood as functions from individuals to propositions. The epistemic and doxastic modal operators are treated as functions from propositions to propositions. The main criticism against this approach is that the concept of a proposition is too-coarse grained; an agent is logically omniscient as it knows/believes all consequences of its knowledge/beliefs [13, 12, 34, 35, 43, 46]. Furthermore, if two sentences denote the same sets of possible worlds, then an agent who knows one is required to know the other. A more fine-grained notion of proposition is required, one that will not compel the agent to knowing all logical consequences of its knowledge.

Alternatively, one may stay within the boundaries of first order logic and adopt a syntactical approach. The propositional attitudes are expressed as predicates in which sentences can occur as terms and these terms are syntactic objects. There are philosophical objections to the syntactical view of propositions as well. At this point, a distinction needs to be made between the syntactical approach stemming mainly from the AI community and Computer Science and the one emanating from Logic and Philosophy. For the discussion here I will refer to the former as syntactic approaches and the latter as axiomatic theories of propositions.

In a syntactic approach such as the one implicitly advocated by Konolige [29], propositions are identified as sentences in some language of representation and they are considered to be syntactic objects. This view seems to provide the fine-grained notion of a proposition required for knowledge representation. However, in the presence of self-reference, it is this

representational view that leads to inconsistency and raises some major philosophical objections. Intuitively, propositions are not the sentences themselves, but rather their semantic content. In other words, when an agent knows a proposition it is not the symbols of the language that it knows, but rather the abstract object that the sentence denotes. As one would expect, human agents do not identify propositions with their syntactic representation, but have an understanding that goes beyond syntax. The problem with the syntactic approach as has been advocated by most researchers in the AI community, originates from the false identification of sentences with propositions. Since sentences pick out propositions uniquely, it is very tempting to regard that there is a one-one correspondence between sentences and propositions. Inevitably however, there are sentences that do not denote any proposition, such as the paradoxical ones. Alternatively, axiomatic theories of propositions do not state what propositions are, but instead they characterise the notion of being a proposition axiomatically. This is the *weak representational* view where although propositions are denoted by sentences, the syntactic category of objects denoting propositions is a subset of the wff of the language. Hence, propositions are those objects that can be characterised as being true or false. In order to avoid the problems of both the aforementioned views a third view of propositions as collections or arrangements of objects and properties in the world has been advocated [52].

### 3 Syntactical Approaches

A rational agent in a multi-agent world needs to be able to reason about the world and what holds true as well as about its own cognitive state and that of the other agents. This involves statements that inevitably make reasonable assertions about beliefs, knowledge and other attitudes. Such statements may require quantification over the objects of the propositional attitudes or that such objects be denoted by terms. Moreover, an agent may have to refer to “oneself” in a number of contexts either directly or indirectly. Consider the following statements:

- 1: John believes things which are false
- 2: John knows all the desires of Carmen
- 3: John knows Carmen has some desires she intends
- 4: Carmen and John share the same desires
- 5: Carmen believes John knows all her beliefs
- 6: Carmen believes John loves Mary

Statements 1-5 require quantification over propositions and (5) in particular involves self-reference. Hence, agent theories need to be sufficiently expressive so as to facilitate quantification over the objects of belief, knowledge and other propositional attitudes as well as the ability to refer to oneself. In classical propositional modal logic modal operators can express all the above attitudes which apply to simple propositions and therefore it is impossible to capture the intuitive meaning of Carmen’s and John’s attitudes. An attempt to formalise them in classical first order logic fails altogether since a predicate of the form  $K_i(A)$  where  $i$  the agent and  $A$  is the object of knowledge is not a wff of predicate calculus. As a result, it is not possible to express propositional attitudes as predicates in classical first order logic. Although first order modal logic is another possible candidate, statements 1-5 cannot be expressed. For instance, (1) requires a variable that ranges over propositions and this is not allowed in this particular framework. A higher order logic such as [39] can be used and (1) would then be formalised as:

$$\exists x_{\langle w,t \rangle} (B_{John}(x) \wedge \neg(x))$$

The variable  $x$  ranges over the type of propositions ( $x_{\langle w,t \rangle}$ ), where propositions are taken to be functions from worlds to truth values. This typed nature of higher order logics does not accord well with human reasoning which is type-free. In addition, higher order logics are considerably more complex than their first order counterparts and they are also computationally intractable.

Alternatively, one may opt for a syntactical approach within first order logic. In such an approach statements about the world can be treated as objects in the world, in other words sentences can be treated as terms in a first order language. This offers several advantages. Firstly, it seems quite natural to be able to represent an agent’s attitudes as sentences in some language. Secondly, since they are formalised within first order logic they may lend themselves directly to implementations [41]. Moreover, they fit well with the representation of intensionality in natural language [4]. The example statements above could be naturally represented as follows:

- 1:  $\exists x B_{John}(x) \wedge F(x)$
- 2:  $\forall x D_{Carmen}(x) \Rightarrow K_{John}(D_{Carmen}(x))$
- 3:  $K_{John}(\exists x D_{Carmen}(x) \wedge I_{Carmen}(x))$
- 4:  $\forall x D_{Carmen}(x) \Leftrightarrow D_{John}(x)$
- 5:  $B_{Carmen}(\forall x B_{Carmen}(x) \Rightarrow K_{John}(B_{Carmen}(x)))$
- 6:  $B_{Carmen}('loves(John, Mary)')$

The reading of the predicate  $K$  for knowledge is intuitively different from the reading of the corresponding modal operator  $K$ . For one, the modal operator takes a sentence as its argument while the predicate takes a term which is the name of the formula in the language. This term denotes a proposition. A formula of the form  $K_i(\phi)$  is interpreted as “agent  $i$  knows that  $\phi$ ”, while  $K_i(\ulcorner \phi \urcorner)$  is read “agent  $i$  knows  $\ulcorner \phi \urcorner$ ”. In the syntactical approach ‘that’-clauses are treated as singular terms which may contain externally quantifiable occurrences of variables. The difference between syntactic theories and modal theories may be considered to be roughly analogous to the difference between direct quotation (“Carmen believes, ‘John loves Mary’”), and indirect quotation (“Carmen believes that John loves Mary”). Any kind of sayable or writable string may be embedded in direct quotation, while only meaningful sentences may be embedded in indirect quotation; “John knows ‘kjhdka sdfas’” is meaningful, while (“John knows that kjhdka sdfas”) is not [8].

Consider sentence 6:  $'loves(John, Mary)'$  is a term denoting a proposition. An external agent required to reason from 6 about Carmen’s beliefs needs to unquote what Carmen believes and reason with the resulting proposition. So any fully expressive logic of propositional attitudes requires terms denoting propositions and a way of unquoting such terms. The way in which this unquoting is achieved is by the introduction of a Truth predicate, which says of its argument that the proposition it denotes is true, so that  $T('loves(John, Mary)')$  asserts the truth of the proposition  $loves(John, Mary)$ :  $T('loves(John, Mary)') \Leftrightarrow loves(John, Mary)$

The expressive power of the syntactical approach facilitates among other things quantification over the objects of the propositional attitudes in an elegant way. This is a highly desirable feature for a theory of reasoning agents; one which is lacking from current approaches [5, 55, 47, 60].

Promising as the syntactical approach may sound, it suffers from a major drawback.

### 3.1 Syntactical Theories and Inconsistency

In the syntactical approach sentences can be treated as terms in a first order language. There are two ways in which this can be achieved:

a) Through arithmetisation (diagonalisation). The theory includes certain fragments of arithmetic and it can talk about its own syntax by identifying syntactic objects with natural numbers. Thus, it can represent its syntactic properties and relations by means of arithmetical predicates. In this way self-referential statements can be constructed by having sentences referring to their own codes.

b) Through a method of quotation. For instance, the term '*loves(john, mary)*' represents the proposition *loves(john, mary)*. The unquoting of the term is achieved through the use of a Truth predicate which says of its argument that the proposition it denotes is true:

$$T('loves(john, mary)') \Leftrightarrow loves(john, mary) \quad (\text{Tarskian Biconditional})$$

However the syntactical approach has one major disadvantage: it is prone to inconsistency. The above schema for unquoting a term in a syntactical theory allows the construction of a sentence which asserts its own falsity: "This sentence is false", known as the *Liar* [3]. The contradiction stems from a combination of the following features of a logical theory  $\mathcal{T}$ :

- (i) construction of statement names as well as self-referential statements in the language
- (ii) acceptance of ordinary logical laws
- (iii) acceptance of the passage from the truth of statement to a statement itself and vice versa

Assume a logical system  $\mathcal{T}$  with a specified syntax and an underlying logic. The small Greek letters ' $\phi$ ', ' $\psi$ '... range over the statements of the logical language  $\mathcal{L}$ , which are also closed under the classical propositional operators  $\wedge, \vee, \neg, \Rightarrow, \Leftrightarrow$ . The customary rules of propositional calculus hold. A formula  $\phi$  with at most one free variable  $x$  is denoted by  $\phi(x)$  and for  $t$  an individual term  $\phi(t)$  is the result of substituting  $t$  for  $x$  in  $\phi$ . Naming in the language is included by allowing each statement  $\phi$  of  $\mathcal{L}$  to have a name, that is with each  $\phi$  is associated a closed term  $\ulcorner \phi \urcorner$ . Then self-referential statements can be constructed as follows: for each formula  $\psi(x)$  a statement  $\phi$  which is equivalent to  $\psi(\ulcorner \phi \urcorner)$  can be constructed. The following axiom schema is accepted for a truth predicate  $T(\phi)$  which is interpreted as " $\phi$  is true":

$$\text{TB } T(\ulcorner \phi \urcorner) \Leftrightarrow \phi \quad \forall \phi \in \mathcal{L}$$

The Diagonalisation Lemma guarantees that there is a  $\phi$  such that  $\phi \Leftrightarrow \neg T(\ulcorner \phi \urcorner)$  [49]. Using the TB-schema and transitivity of equivalence  $T(\ulcorner \phi \urcorner) \Leftrightarrow \neg T(\ulcorner \phi \urcorner)$  is obtained. Let  $a = T(\ulcorner \phi \urcorner)$ . Consequently, both  $(a \Rightarrow \neg a)$  and  $(\neg a \Rightarrow a)$  are provable. From ordinary propositional reasoning  $(a \Rightarrow \neg a) \Rightarrow \neg a$  is derived, and hence  $\neg a$  follows by modus ponens. But then  $a$  follows in the same way from the other implication, a contradiction, and thus, the theory  $\mathcal{T}$  is inconsistent.

A similar assertion can be made involving knowledge [40]. Hence, it is very difficult to provide a suitable axiomatisation for the intentional notions without introducing inconsistency. Montague [39] proved negative results regarding the consistency of first order syntactical theories for knowledge. Assuming a first order syntactical treatment of knowledge  $\Gamma$  and the following axiom schemas:

- i)  $K(\ulcorner \phi \urcorner) \Rightarrow \phi$
- ii)  $K(\ulcorner K(\ulcorner \phi \urcorner) \Rightarrow \phi \urcorner)$
- iii) if  $\vdash \phi$  then  $\vdash K(\ulcorner \phi \urcorner)$

$$\text{iv) } K(\ulcorner \phi \Rightarrow \psi \urcorner) \Rightarrow (K(\ulcorner \phi \urcorner) \Rightarrow K(\ulcorner \psi \urcorner))$$

then  $\Gamma$  is inconsistent [39].

Thomason [51], inspired by Löb's [49] version of the Hilbert-Bernays *Derivability Conditions*, extended this result for the weaker notion of belief by replacing i) with:

$$\text{v) } B(\ulcorner \phi \urcorner) \Rightarrow B(\ulcorner B(\ulcorner \phi \urcorner) \urcorner).$$

### 3.2 Avoiding the Inconsistency

In order to retain consistency, three kinds of restrictions may be considered in a syntactical system: *i)* restriction of syntax, *ii)* restriction of logic and finally *iii)* restriction of basic principles.

The first serious attempt to avoid the paradox of the Liar sentence was that of Tarski [50] via a hierarchy of languages. The escape route taken is that of restricting the syntax (*i*). If I begin with an “object language”  $L$ , I can use an enriched “meta-language”  $L'$  with a predicate  $True_L$  that can be applied to the true sentences of the original language. Since the truth predicate applies to all and only the true sentences of  $L$  it will allow me to say (in  $L'$ ) many of the things I wanted to say using the intuitive notion of truth at least for the original object language. Indeed, this use of an object language and a meta-language avoids the paradox. But the same predicate  $True_L$  cannot be used to refer to sentences of the same meta-language. For this reason a new meta-language  $L''$  that stands in the same relation to  $L'$ , as  $L'$  stood to  $L$ , has to be defined. This hierarchy of languages however, does not seem to accord well with our intuitions. Human languages do not have distinct levels each with its own truth predicate. Instead there is only one word “true” which applies to all sentences. Even if human languages are left aside, and the assertions and reasoning of artificial agents are considered, it is still difficult to have knowledge of the level of the agent's utterances. Yet another disadvantage is that if two statements are on the same level, neither can talk about the truth or falsity of the other. Moreover, as Kripke [32] points out, it is impossible to assert that all the statements in such a sequence of languages are true; this would require a meta-language of transfinite level above all the other languages. Finally, by employing such a hierarchy of languages the ability to express self-referential statements is lost altogether.

In an attempt to overcome the problems generated by the Tarskian account for truth, Kripke [32] developed a theory that allowed circular reference and contained its own truth predicate. His solution involves restrictions of the logic (*ii*). He developed a method for approximating the usually unattainable idea of finding an interpretation for a logical language  $L$ , in which the extension of the truth predicate  $T$  coincides with the set of all sentences of  $L$  true on that interpretation. This method is based on an iterative revision process in which repeated adjustments of the extension of  $T$  are required. Each progressive adjustment consists of making the extension of  $T$  equal to the set of sentences that have just been determined as true. Kripke used the Kleene strong three-valued logic, although other valuation schemes can be used as well. The Kripkean account of truth permits truth-value gaps; sentences need not be true or false. Thus, paradoxical sentences are neither true nor false, but they are rendered “unknown”. Those sentences that are assigned a truth value after the revision process are said to be “grounded” and the rest are said to be “ungrounded”. Kripke proves two fundamental results about the revision process. Firstly, the revision process is monotonic, and secondly it eventually terminates: a saturation point is reached where no new sentences are added to the extension or anti-extension of the truth predicate.

Motivated by the work of Kripke, Gupta [21] and Herzberger [24] developed alternative

approaches towards a theory of Truth. Rather than adopting the Kleene strong three-valued logic for the revision process, they remained within classical two-valued logic. They imposed restrictions on the basic principles governing the truth predicate, in particular they rejected the principle of bivalence  $T(\phi) \vee T(\neg\phi)$ . Their method is based on an iterative revision process as well, but assigns truth values to a larger set of sentences than Kripke’s theory does. For instance, the Truth Teller (“This sentence is true”) that is ungrounded in Kripke’s account, can take a truth value in the Gupta-Herzberger semantic theory of truth. Following Gupta and Herzberger, Asher and Kamp [2] investigated theories which could be used as the basis for logics of knowledge and belief. Their work combines the Gupta-Herzberger iterative revision process with a possible worlds semantics. Instead of having a truth predicate, they consider a belief predicate  $B$  which is to be taken as true of all and only those sentences which are believed to be true by some fixed agent  $a$ . The sentences that are believed are all those that are true in all the belief-possible worlds according to the agent. The belief predicate  $B$ , as the truth predicate in the Gupta-Herzberger semantic theory of truth, has an extension now at every possible world  $w$ . This extension should coincide with the set of sentences true in all of  $w$ ’s alternative possible worlds. The iterative revision process is called to adjust the extension of  $B$  at  $w$  when the two sets of sentences are not the same. Although they do not offer a single system as an epistemic or doxastic logic, they provide useful results regarding coherency in different models  $\mathcal{M}$  of a first order language  $\mathcal{L}$  depending on the accessibility relation  $\mathcal{R}$  over possible worlds, the initial extension of the model ( $\|B\|_{\mathcal{M},w}^0$ ) and the kinds of self-reference that are allowed in  $\mathcal{M}$  (quantificational or designative).

### 3.3 Related Work in Artificial Intelligence

Syntactical theories seem to be an attractive way of formalising the intentional notions and the properties of artificial agents. Such agents should be able to make assertions about their knowledge and beliefs as well as about the other agents’ knowledge and beliefs. Within Artificial Intelligence, among the first to propose the use of a syntactical approach was McCarthy [37] with a first order language that includes concepts (terms) that are the names of the terms and the formulae of the language.

An alternative approach was advocated by Konolige [28], with a hierarchical meta-language framework for reasoning about an agent’s beliefs, knowledge, and actions. In Konolige’s formalism the agent’s beliefs are represented as sentences in an object language OL which is a first order logic with equality. Then the propositional attitude of belief is expressed as a predicate in the meta-language ML. More specifically, Konolige introduces a meta-language function  $th$  which takes as argument an agent and returns the agent’s theory of the world, i.e. a set of object-language formulae. Three predicates are introduced in the meta-language:

- i) TRUE( $f$ )*, expresses that the object language formula is true in a specific world
- ii) FACT( $t, g$ )*, expresses that  $f$  is one of the base formulae of theory  $t$
- iii) PR( $t, f$ )*, means that  $f$  is provable in theory  $t$ , and it can be given an axiomatisation as required

Knowledge which is defined as true belief, is introduced in the meta-language as a predicate. The formalism also includes the situation calculus, a standard name function, and a denotation function to overcome some of the object/meta-language problems. Nested beliefs can be expressed by extending the two-language hierarchy to a three-language one. This approach suffers unfortunately from several drawbacks:

- a) a three-language hierarchy requires complex technical machinery for linking the lan-



guages

- b) the agent's nested beliefs are limited to two levels
- c) only one agent is allowed to reason about the other agent's beliefs
- d) computationally the approach seems to be inefficient and expensive due to the meta-level factor
- e) some of the axioms of belief commit the agent to logical omniscience

Perlis [44, 45] also investigated the possibility of a first order logic of truth and knowledge. Perlis starts by weakening the Tarskian Biconditional  $T(A) \Leftrightarrow A$  by replacing it with the schema  $T(A) \Leftrightarrow A^*$  based on the work of Gilmore [19].  $A^*$  is the result of replacing in a formula  $A$  each connective occurrence of the form  $\neg T(B)$  by  $T(\neg B)$ . Although the law of excluded middle holds, the law of bivalence does not hold ( $T(A) \vee T(\neg A)$ ). In this way paradoxical sentences are dealt with since  $T(A)$  and  $T(\neg A)$  are both allowed at the same time. Perlis also considers a syntactic modality for belief with a K4 axiomatisation and then knowledge is defined as true belief.

Turner [53] built an axiomatic theory of propositions, truth and modality based upon the ideas of Kripke [32], Gupta [21] and Herzberger [24]. He advocated a weak representationalist view that supports a fine-grained notion of proposition. Only a subclass of sentences denote propositions, and these are the objects that can be true or false. He investigated various logics of stable truth and how these logics can be integrated with standard modal systems to produce useful systems for Knowledge Representation. Following the work of Turner, Davies [6] developed a self-referential first order language with a truth predicate whose terms are those of the Lambda Calculus. In order to achieve consistency and avoid the Tarskian paradox propositions are confined only to those sentences that are true or false:

$$P(A) =_{def} T(A) \vee T(\neg A)$$

The proposed logic includes an axiomatisation of truth and belief, while knowledge is introduced as true belief.

More recently Morreau and Kraus [41] advocated a syntactical treatment for the propositional attitudes and extended previous results of des Rivieres and Levesque [10]. In their work des Rivieres and Levesque investigated the consistency of theories in which the axiom schemas for knowledge do not range over all sentences of the language, but only over a subset of them called regular sentences. While this result is very useful, certain things cannot be expressed with regular formulae. For instance, quantification over the objects of the propositional attitudes cannot be instantiated in a knowledge base without introducing inconsistency. Morreau and Kraus extend the applicability of the axiom schemes of knowledge (or belief) to what they call RPQ-formulae (regular formulae with propositional quantification). RPQ formulae include the regular formulae of des Rivieres and Levesque, but in addition quantification over the propositional attitudes is permitted through the use of two additional predicates  $P$  (picks out the Gödel numbers of sentences of the original language  $L_a$ ) and  $T$  (picks out the true sentences of  $L_a$ ). These two new predicates take as arguments new variables  $X, Y, \dots$  which instantiate with Gödel numbers of sentences. In this way the RPQ formulae are more expressive than regular formulae since they allow quantification over the propositional objects. Although a Truth predicate is introduced, no semantics or axiomatisation is given. The result of the use of RPQ formulae is that formulae that quantify over the objects of propositional attitudes can be included in a knowledge base without introducing inconsistency; it is the scope of the axiom schemes that is being restricted and not the use of other formulae in the knowledge base which can contain other formulae than RPQ.

## 4 Formal Framework

In the following sections the details of the logical language and the basic building blocks of the theory are spelt out. Proofs are provided in Appendix A.1.

### 4.1 Logical Language

To facilitate principles like the Tarskian Biconditional some statement-naming and/or abstraction devices are needed. For reasons of generality and simplicity an underlying first order language  $\mathcal{L}_0$  and a first order theory  $\mathsf{T}_0$  are assumed. Initially, assume that the language  $\mathcal{L}_0$  contains a constant symbol  $\bar{0}$  also written  $0$ , a binary operation symbol  $d_r$  and unary operation symbols  $q_1$  and  $q_2$ . Moreover, the underlying first order theory  $\mathsf{T}_0$  has the following formulae provable:

- (i)  $d_r(x, y) \neq 0$
- (ii)  $q_1(x, y) = x \wedge q_2(x, y) = y$

Consequently,  $d_r( , )$  acts as a pairing operation from  $M^2 \rightarrow M - \{0\}$ , for which  $q_1$  and  $q_2$  are the corresponding projection operations. In  $\mathsf{T}_0$  the structure of natural numbers can be represented by defining  $x' = (x, \bar{0})$ . Thus,  $q_1$  acts as the predecessor operation and from the axioms of  $\mathsf{T}_0$  I derive:

- (iii)  $x' \neq 0$
- (iv)  $x' = y' \rightarrow x = y$

Tuples  $(t_1, \dots, t_k)$  can be introduced recursively by  $(t_1) = t$  and  $(t_1, \dots, t_{k+1}) = ((t_1, \dots, t_k), t_{k+1})$ . Moreover, there are corresponding projection operations  $q_i^k (1 \leq i \leq k)$  satisfying  $q_i^k(t_1, \dots, t_k) = t_i$  in  $\mathsf{T}_0$ .

The language of the theory  $\mathcal{L}$  is based on [4] and it is a first order language which in addition to the apparatus of  $\mathcal{L}_0$  contains:

- (i) a set of connective symbols  $\{\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow\}$
- (ii) a set of quantifiers  $\{\forall, \exists\}$
- (iii) a set of punctuation symbols  $\{(, ), , \}$
- (iv) a predicate  $=$  for expressing equality
- (v) a set of variables  $V$
- (vi) a set of relation symbols  $P_{k \geq 0}$

The terms and wff of  $\mathcal{L}$  are defined by (simultaneous) recursion as follows:

- (i) every variable, constant or term of  $\mathcal{L}_0$  is a term of  $\mathcal{L}$
- (ii) if  $P$  is a predicate letter and  $t_1, \dots, t_k$  are terms, then  $P(t_1, \dots, t_k)$  is an atomic wff
- (iii) if  $t_1$  and  $t_2$  are terms, then  $t_1 = t_2$  is a wff
- (iv) if  $A$  and  $B$  are wff, so are  $\neg A$ ,  $A \wedge B$ ,  $A \vee B$ ,  $A \Rightarrow B$  and  $A \Leftrightarrow B$
- (v) if  $A$  is a wff and  $x_1, \dots, x_k, y_1, \dots, y_n$  include all the variables of  $A$ , then  $\hat{x}_1, \dots, \hat{x}_k[A]$  is a term in which  $x_1, \dots, x_k$  are bound and  $y_1, \dots, y_n$  are free.
- (vi) every wff of  $\mathcal{L}_0$  is a wff of  $\mathcal{L}$
- (vii) if  $x$  is a variable and  $A$  is a wff, then  $\forall x A$  and  $\exists x A$  are wff

The existential quantifier can be defined in terms of the negation symbol and the universal quantifier as  $\exists x A =_{def} \neg \forall x \neg A$ . If a variable occurs inside the scope of a quantifier the occurrence of the variable is said to be bound; otherwise the occurrence of the variable is said to be free.

Following Feferman [17] the terms of  $\mathcal{L}$  could be introduced as terms of  $\mathcal{L}_0$ . For a wff  $A$ , using the apparatus of  $\mathcal{L}_0$  let  $r_A$  be the Gödel number of  $A$ . If  $A$  has free variables among

$x_1, \dots, x_k, y_1, \dots, y_n$  then  $(r_a, y_1, \dots, y_n)$  serves as an operation in  $\mathcal{L}_0$  which “abstracts”  $x_1, \dots, x_k$  treating  $y_1, \dots, y_n$  as parameters. I can then define:

$$\Delta_k \quad \widehat{x}_1, \dots, \widehat{x}_k[A(x_1, \dots, x_k, y_1, \dots, y_n)] = (r_A, y_1, \dots, y_n)$$

Hence,  $(r_A, y_1, \dots, y_n)$  is the term representing the formula  $A(x_1, \dots, x_k, y_1, \dots, y_n)$  and by definition is a term of  $\mathcal{L}$ . Here, I will simply write  $A$  instead of  $(r_A, y_1, \dots, y_n)$  where no confusion can arise. For  $k = 0$ :

$$\Delta_0 \quad [A(y_1, \dots, y_n)] = (r_A, y_1, \dots, y_n)$$

which is identified with  $r_A$  when  $n = 0$ . The purpose of the denotational device  $\Delta_k$  is in connection with the truth axiom. For each  $k$  a  $(k + 1)$ -placed predicate symbol  $T_k$  is introduced where  $T_k(x_1, \dots, x_k, A)$  is read “ $(x_1, \dots, x_k)$  satisfies  $A$ ”. For the special case  $k = 0$ , a truth predicate  $T$  is introduced which operates on terms, i.e.  $T_0 = T$ .

(viii) if  $t$  is a term, then  $T(t)$  is a wff

A false predicate is then defined as  $F(t) =_{def} T(\neg t)$ .

## 4.2 Model Theory

A model for the logical language  $\mathcal{L}$  is a tuple  $M = \langle U, \pi, \mathcal{T}, \mathcal{F} \rangle$  where:

-  $U$  is the universe of discourse with the following restrictions:

(i) the domain of individuals is constant, and

(ii) the domain  $U$  is cartesian closed (i.e. is closed under the formation of tuples; if  $d_1, d_2, \dots, d_n \in U$  then  $(d_1, d_2, \dots, d_n) \in U$ ).

-  $\pi$  is used to determine the truth values of the atomic formulae of the language, apart from the truth and the false predicate.

-  $\mathcal{T}$  is the extension of the truth predicate, that is  $\mathcal{T} : U \rightarrow \{0, 1\}$ .

-  $\mathcal{F}$  is the extension of the false predicate, that is  $\mathcal{F} : U \rightarrow \{0, 1\}$ .

The additional condition  $\mathcal{T} \cap \mathcal{F} = \emptyset$  is imposed on the extensions of  $\mathcal{T}$  and  $\mathcal{F}$  so that no term can be both true and false.  $\| \cdot \|$  is defined as follows:

$$\| x \|_v = v(x)$$

$$\| 0 \|_v = 0 \in U$$

$$\| (t_1, t_2) \|_v = (\| t_1 \|_v, \| t_2 \|_v)$$

The support relations where  $v$  is a mapping of variables into  $U$ , are as follows:

$$M \models_v P(t_1, \dots, t_k) \text{ iff } \langle v(t_1), \dots, v(t_k) \rangle \in \pi(P^k)$$

$$M \models_v \neg A \text{ iff } M \not\models_v A$$

$$M \models_v T(t) \text{ iff } \mathcal{T}(\| t \|_v) = 1$$

$$M \models_v F(t) \text{ iff } \mathcal{F}(\| t \|_v) = 1$$

$$M \models_v A \wedge B \text{ iff } M \models_v A \text{ and } M \models_v B$$

$$M \models_v \forall x A \text{ iff for all } d \text{ in } U, M \models_{v[d/x]} A$$

Similarly for the other connectives and quantifier.

## 4.3 Proof Theory

The formal system of first order predicate calculus with equality is adopted. For any wff  $A$ ,  $B$  and  $C$  of  $\mathcal{L}$ :

$$A \Rightarrow (B \Rightarrow A)$$

$$(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$$

$$(\neg B \Rightarrow \neg A) \Rightarrow ((\neg B \Rightarrow A) \Rightarrow B)$$

$$\forall x A(x) \Rightarrow A(t) \text{ where } t \text{ is a term free of } x \text{ in } A(x)$$

$$(\forall x (A \Rightarrow B)) \Rightarrow (A \Rightarrow \forall x B) \text{ where } A \text{ contains no free occurrences of } x$$

$A, A \Rightarrow B \vdash B$

$A \vdash \forall x A$

I write  $\text{FOPC} \vdash A$  if  $A$  is provable in this system.

## 5 Stable Truth

The intuitions for the truth predicate dictate that whatever is asserted to be true, must be so. Thus, the appropriate axiom schema (Tarskian Biconditional) is as follows:

$T_k B.$   $T_k(\widehat{x}_1, \dots, \widehat{x}_k, [A(x_1, \dots, x_k, y_1, \dots, y_n)]) \Leftrightarrow A(x_1, \dots, x_k, y_1, \dots, y_n)$

For  $k = 0$  this reduces to:

$T_0 B.$   $T_0([A(y_1, \dots, y_n)]) \Leftrightarrow A(y_1, \dots, y_n)$

and in particular, for  $n = 0$  it reduces to:

$T B.$   $T([A]) \Leftrightarrow A$

However, the inclusion of such an axiom schema in the logic leads to inconsistency. In order to maintain consistency something must be given up. In the Gupta-Herzberger [21, 24] semantic theory the escape route to the semantic paradoxes is to discard the principle of bivalence:

Biv.  $T(A) \vee T(\neg A)$

In other words, not all sentences in the language denote propositions; some sentences are paradoxical. The intuitions underlying the Gupta-Herzberger [21, 24] semantic theory are based on the idea of a semi-inductive iterative revision process. This revision process starts with simple statements that do not contain the word true and they are assigned a truth value according to the empirical facts. As the process continues, more and more statements involving complex assertions about truth and falsity are assigned a truth value.

Given the model  $M$  for  $\mathcal{L}$ , as defined above, define  $M' = \langle U, \pi, \mathcal{T}', \mathcal{F}' \rangle$  the Tarskian revision of  $M$  such that:

$$\mathcal{T}'(\| t \|_v) = \begin{cases} 1 & \text{if } M \models_v A(y_1, \dots, y_n) \text{ where } t = (r_A, y_1, \dots, y_n) \\ \mathcal{T}(\| t \|_v) & \text{otherwise} \end{cases}$$

$$\mathcal{F}'(\| t \|_v) = \begin{cases} 1 & \text{if } M \models_v \neg A(y_1, \dots, y_n) \text{ where } t = (r_A, y_1, \dots, y_n) \\ \mathcal{F}(\| t \|_v) & \text{otherwise} \end{cases}$$

where  $v$  is an assignment of elements of  $U$  to variables. Only the wff change since for the other elements such as constants and variables the  $\mathcal{T}$  and  $\mathcal{F}$  do not change. Starting from a model  $M$  with arbitrary extensions of  $\mathcal{T}$  and  $\mathcal{F}$  and using the above revision step, a sequence of  $\mathcal{T}(n)$  and  $\mathcal{F}(n)$  for  $n \geq 0$  can be defined:

i)  $\mathcal{T}(0) = \mathcal{T}$

$\mathcal{F}(0) = \mathcal{F}$

ii)  $\mathcal{T}(n+1) = \mathcal{T}(n')$

$\mathcal{F}(n+1) = \mathcal{F}(n')$

iii) for a limit ordinal  $k$  define

$\mathcal{T}(k)(d) = 1$  iff  $\exists j(j < k) \forall h(j \leq h < k) (\mathcal{T}(h)(d) = 1)$

$\mathcal{F}(k)(d) = 1$  iff  $\exists j(j < k) \forall h(j \leq h < k) (\mathcal{F}(h)(d) = 1)$

Under this revision process the notion of stability can be defined as:

- an element  $d$  of  $U$  is positively stable iff  $\exists j \forall k \geq j \mathcal{T}(k)(d) = 1$

- an element  $d$  of  $U$  is negatively stable iff  $\exists j \forall k \geq j \mathcal{F}(k)(d) = 1$
- $d$  is stable iff it is positively or negatively stable
- $d$  is positively stable from  $k$  iff  $\forall j \geq k \mathcal{T}(j)(d) = 1$
- $d$  is negatively stable from  $k$  iff  $\forall j \geq k \mathcal{F}(j)(d) = 1$

The notion of a stabilisation ordinal is central to the Gupta-Herzberger approach. A stabilisation ordinal is that point in the revision process at which no more objects will become (stably) true or false, that is a saturation point has been reached where nothing more can be said about truth and falsity. An ordinal  $\sigma$  is a stabilisation ordinal iff:

- i)  $\forall d \in U$ ,  $d$  is positively stable iff  $\mathcal{T}(\sigma)(d) = 1$   
 $\forall d \in U$ ,  $d$  is negatively stable iff  $\mathcal{F}(\sigma)(d) = 1$
- ii)  $\forall d \in U$ ,  $d$  is positively (negatively) stable implies that  $d$  is positively (negatively) stable from  $\sigma$ .

**Theorem 1** ([24]) *There exists a stabilisation ordinal.*

There are two notions of validity that are of interest. The first is generated by the class of sentences that are true at all stabilisation models, while the second by the class of sentences that are stably true. These are defined as follows:

- Definition 2** a) a wff  $A$  is safe iff  $A$  is valid at every stabilisation ordinal.  
b) a wff  $A$  is stably true iff  $T(A)$  is safe and it is stably false iff  $F(A)$  is safe.

## 5.1 Basic Logic of Truth

Consistent logics of truth for knowledge representation were investigated by Turner [53]. The standard modal logic D will serve as the basic logic of truth here. The following naming convention is adopted: the name of the axiom will be indicated by the capital letter T (for truth) followed by a dash and the name of the axiom as is usually referred to in the literature. Small modifications will be made and where the classical axioms S5 and S4 are used, I will write them as 5 and 4 instead. The necessitation rule will be indicated by an N and the Barcan formula by a B.

- T-K.  $T(A \Rightarrow B) \Rightarrow (T(A) \Rightarrow T(B))$
- T-D.  $T(A) \Rightarrow \neg T(\neg A)$
- T-B.  $\forall x T(A) \Rightarrow T(\forall x A)$
- T-N. If  $\vdash A$  then  $\vdash T(A)$

**Theorem 3** ([53]) *D is a consistent logic of truth and all the theorems of D are stably true.*

Additional axioms can be considered in order to strengthen the logic such as the standard axioms T-T, T-4, T-5:

- T-T.  $T(A) \Rightarrow A$
- T-4.  $T(A) \Rightarrow T(T(A))$
- T-5.  $\neg T(A) \Rightarrow T(\neg T(A))$

However, adding the T-T axiom to the D logic of truth results in inconsistency.

**Theorem 4** *The logic of Truth T is inconsistent.*

Turner [53] proves that although T-T and T-4 are safe they are not stably true, whereas T-5 is not even safe. Thus if consistency is to be retained, none of these axioms can be added to the logic without making additional compromises.

## 5.2 Restricted Logics of Truth

The logic D described in the previous section utilises the full rule of necessitation which allows the inference from any valid wff  $A$  to the wff  $T(A)$ . Although the T-T and T-4 axioms cannot be allowed in combination with the full rule of necessitation since they do not preserve stability, nevertheless they are safe and can be added to the logic of truth provided that the rule of necessitation is restricted. This is achieved by prohibiting the application of the rule to any theorems derived using T-T or T-4.

Let  $T[X]$  be the logic which consists of the T-K, T-D, and the T-B axiom (T is not to be confused for the standard system T but instead here is used to denote the logic of the truth predicate) and  $X$  be a subset of  $\{T-T, T-4\}$ , the safe axioms of truth.

**Theorem 5** *The T-RN1 rule can be consistently added to the  $T[X]$  logics of truth:*  
*T-RN1. If  $\vdash_{FOPC} A$  then  $\vdash T(A)$*

Moreover, the stronger T-N rule can be added to the family of logics  $T[X]$ :

**Theorem 6** *The T-RN rule can be consistently added to the  $T[X]$  logics of truth:*  
*T-RN. If  $D \vdash A$  then  $\vdash T(A)$*

The rule is now called T-RN to indicate that although it is used in the family of logics  $T[X]$  its application is restricted. The application of T-RN is allowed on a wider set of formulae than that of T-RN1 (theorems of FOPC), formulae that have been derived with the T-K, T-D and the T-B axioms. From this point onwards the family of logics  $T[X]$  will be understood as including this weakened rule.

**Theorem 7** *The logics  $T[X]$  where  $X$  is  $\{T-T\}$  or  $\{T-T, T-4\}$  are consistent logics of truth.*

## 6 Modalities

The next step in the construction of the logics is to incorporate modal operators for the intentional notions. A modal operator for knowledge and one for belief are added, which for the time being are independent of one another. The logical language  $\mathcal{L}$  is extended to  $\mathcal{L}_1$  with the following two clauses:

- (i) if  $A$  is a wff, then  $Know_i(A)$  is a wff
- (ii) if  $A$  is a wff, then  $Bel_i(A)$  is a wff

The index  $i$  identifies the individual agent. The classical possible worlds semantics will be employed to interpret the modal operators [26]. In line with this, the model theory of stable truth needs to be reformulated as well. A model for the logical language  $\mathcal{L}_1$  is a tuple  $M = \langle W, \mathcal{K}_i, \mathcal{B}_i, U, \pi, \mathcal{T}, \mathcal{F} \rangle$  where:

- $W$  is a set of possible worlds.
- $\mathcal{K}_i$  are the knowledge accessibility relations for each agent  $i$  of the domain.
- $\mathcal{B}_i$  are the belief accessibility relations for each agent  $i$ .
- $U$  is the domain of discourse with the additional restriction that each individual constant is a rigid designator, i.e. denotes the same individual in all possible worlds.
- $\pi$  is used to determine the truth values of the atomic formulae of the language, apart from the truth and the false predicate; it is a function from predicate letters and possible worlds into relations on  $U$  of the appropriate arity.

- $\mathcal{T}$  is the extension of the truth predicate, that is  $\mathcal{T} : U \times W \rightarrow \{0, 1\}$ .
- $\mathcal{F}$  is the extension of the false predicate, that is  $\mathcal{F} : U \times W \rightarrow \{0, 1\}$ .

The conditions supporting the truth in such models are defined by the following clauses where  $v$  is a mapping of variables into  $U$  and  $w, w' \in W$ :

$$\begin{aligned}
M \models_{v,w} P(t_1, \dots, t_k) &\text{ iff } \langle v(t_1), \dots, v(t_k) \rangle \in \pi(P^k, w) \\
M \models_{v,w} \neg A &\text{ iff } M \not\models_{v,w} A \\
M \models_{v,w} Know_i(A) &\text{ iff for all } w' \text{ such that } \mathcal{K}_i(w, w') \text{ } M_{v,w'} \models A \\
M \models_{v,w} Bel_i(A) &\text{ iff for all } w' \text{ such that } \mathcal{B}_i(w, w') \text{ } M_{v,w'} \models A \\
M \models_{v,w} T(t) &\text{ iff } \mathcal{T}(\| t \|_{v,w}) = 1 \\
M \models_{v,w} F(t) &\text{ iff } \mathcal{F}(\| t \|_{v,w}) = 1 \\
M \models_{v,w} A \wedge B &\text{ iff } M \models_{v,w} A \text{ and } M \models_{v,w} B \\
M \models_{v,w} \forall x A &\text{ iff for all } d \text{ in } U, M \models_{v[d/x],w} A
\end{aligned}$$

Similarly for the other connectives and quantifier.

The Tarskian revision of the model  $M$  is defined as  $M' = \langle W, \mathcal{K}_i, \mathcal{B}_i, U, \mathcal{T}', \mathcal{F}' \rangle$  where  $\mathcal{T}$  and  $\mathcal{F}$  are now relativised to possible worlds:

$$\mathcal{T}'(\| t \|_{v,w}) = \begin{cases} 1 & \text{if } M \models_{v,w} A(y_1, \dots, y_n) \text{ where } t = (r_A, y_1, \dots, y_n) \\ \mathcal{T}(\| t \|_{v,w}) & \text{otherwise} \end{cases}$$

$$\mathcal{F}'(\| t \|_{v,w}) = \begin{cases} 1 & \text{if } M \models_{v,w} \neg A(y_1, \dots, y_n) \text{ where } t = (r_A, y_1, \dots, y_n) \\ \mathcal{F}(\| t \|_{v,w}) & \text{otherwise} \end{cases}$$

The interesting wff are those that are valid in stabilised models which are called  $\Gamma\Delta$ -models.  $\Gamma$  indicates the model for knowledge which can be any of the following:

- D-, the knowledge accessibility relation is serial;
- T-, the knowledge accessibility relation is reflexive and serial;
- S4-, the knowledge accessibility relation is transitive, serial and reflexive;
- S5-, the knowledge accessibility relation is symmetric, transitive and reflexive.

$\Delta$  indicates the model for belief which can be any of the following:

- D-, the belief accessibility relation is serial;
- KD4-, the belief accessibility relation is transitive and serial;
- KD45-, the belief accessibility relation is Euclidean, transitive and serial.

The notion of safeness and stability can be relativised to the class of  $\Gamma\Delta$ -models under consideration:

**Definition 8** A wff  $A$  is  $\Gamma\Delta$ -safe iff it is valid at all stabilisation models for all initial  $\Gamma\Delta$ -models  $M$ .

**Definition 9** A wff  $A$  is  $\Gamma\Delta$ -positively stable if  $T(A)$  is  $\Gamma\Delta$ -safe and it is  $\Gamma\Delta$ -negatively stable if  $F(A)$  is  $\Gamma\Delta$ -safe.

## 6.1 Epistemic and Doxastic Axiomatizations

The logic of the epistemic operator for knowledge *Know* can be any of the standard modal logics D, T, S4, or S5 with full necessitation. The full rule of necessitation is allowed for the modal operator in contrast with the truth axiomatisation. The difference is that the modal operator does not use a term as its argument (which can be a quoted sentence), but a simple

sentence. From this point onwards the following naming conventions for the various axiom schemas will be adopted:

(i) the names of axioms and rule for the modal operator  $Know_i$  will be indicated by a small letter 'k', followed by a dash and the name of the axiom or rule;

(ii) the names of axioms and rule for the modal operator  $Bel_i$  will be indicated by a small letter 'b', followed by a dash and the name of the axiom or rule.

The strongest axiomatisation for knowledge and belief is the S5 and KD45 systems respectively:

### Knowledge

k-K.  $Know_i(A \Rightarrow B) \Rightarrow (Know_i(A) \Rightarrow Know_i(B))$

k-D.  $Know_i(A) \Rightarrow \neg Know_i(\neg A)$

k-T.  $Know_i(A) \Rightarrow A$

k-4.  $Know_i(A) \Rightarrow Know_i(Know_i(A))$

k-5.  $\neg Know_i(A) \Rightarrow Know_i(\neg Know_i(A))$

k-B.  $\forall x Know_i(A) \Rightarrow Know_i(\forall x A)$

k-N. If  $\vdash A$  then  $\vdash Know_i(A)$

### Belief

b-K.  $Bel_i(\phi \Rightarrow \psi) \Rightarrow (Bel_i(\phi) \Rightarrow Bel_i(\psi))$

b-D.  $Bel_i(\phi) \Rightarrow \neg Bel_i(\neg \phi)$

b-4.  $Bel_i(A) \Rightarrow Bel_i(Bel_i(A))$

b-5.  $\neg Bel_i(A) \Rightarrow Bel_i(\neg Bel_i(A))$

b-B.  $\forall x Bel_i(A) \Rightarrow Bel_i(\forall x A)$

b-N. If  $\vdash A$  then  $\vdash Bel_i(A)$

The T axiom is not part of the axiom system for belief. It seems intuitively correct to allow agents to have false beliefs at times. The logic consisting of the D logic for the epistemic operator for knowledge, and the D logic for the doxastic operator of belief will from now on be called  $KB^m$ . The family of logics  $KB^m[X,Y]$  is obtained, where X can be {k-T}, {k-T, k-4} or {k-T, k-4, k-5}, and Y {b-4} or {b-4, b-5}.

## 6.2 Logics for Modalities and Predicate Truth

The first system that can be obtained in which the epistemic and doxastic modal logics are combined with predicate truth is the system in which the strongest axiom is D for all three notions. This system will be called  $TKB^m$ :

### Truth

T-T.  $T(A \Rightarrow B) \Rightarrow (T(A) \Rightarrow T(B))$

T-D.  $T(A) \Rightarrow \neg T(\neg A)$

T-B.  $\forall x T(A) \Rightarrow T(\forall x A)$

T-N. If  $\vdash A$  then  $\vdash T(A)$

### Knowledge

k-K.  $Know_i(A \Rightarrow B) \Rightarrow (Know_i(A) \Rightarrow Know_i(B))$

k-D.  $Know_i(A) \Rightarrow \neg Know_i(\neg A)$

k-B.  $\forall x Know_i(A) \Rightarrow Know_i(\forall x A)$

k-N. If  $\vdash A$  then  $\vdash Know_i(A)$

### Belief

b-K.  $Bel_i(\phi \Rightarrow \psi) \Rightarrow (Bel_i(\phi) \Rightarrow Bel_i(\psi))$

b-D.  $Bel_i(\phi) \Rightarrow \neg Bel_i(\neg \phi)$



- b-B.  $\forall x Bel_i(A) \Rightarrow Bel_i(\forall x A)$
- b-N. If  $\vdash A$  then  $\vdash Bel_i(A)$

**Theorem 10** *If  $TKB^m \vdash A$  then  $A$  is  $\Gamma\Delta$ -stably true where  $\Gamma$  is  $D$  and  $\Delta$  is  $D$  as well.*

There are three possible directions in which the logic can be extended: (i) by adding stronger axioms for the epistemic operator for knowledge; (ii) stronger axioms for the doxastic operator for belief; or finally (iii) stronger axioms for truth. For the first two directions the standard modal axioms k-T, k-4, k-5 for knowledge and b-4 and b-5 for belief respectively, are considered.

**Theorem 11** *a) The axioms k-T, k-4, k-5, for the knowledge operator are  $\Gamma$ -safe where  $\Gamma$  are T-, S4-, S5- models respectively.  
b) The axioms b-4, b-5, for the belief operator are  $\Delta$ -safe where  $\Delta$  are KD4-, KD45- models respectively.*

Let Y be any subset of the modal axioms for knowledge  $Y = \{k-T, k-4, k-5\}$  and Z be any subset of the modal axioms for belief  $\{b-4, b-5\}$ , then  $TKB^m[Y, Z]$  is the  $TKB^m$  logic plus the subsets Y and Z together with the full rules of necessitation:

- T-FN. if  $TKB^m[Y, Z] \vdash A$  then  $TKB^m[Y, Z] \vdash T(A)$
- k-FN. if  $TKB^m[Y, Z] \vdash A$  then  $TKB^m[Y, Z] \vdash Know_i(A)$
- b-FN. if  $TKB^m[Y, Z] \vdash A$  then  $TKB^m[Y, Z] \vdash Bel_i(A)$

**Theorem 12** *If  $TKB^m[Y, Z] \vdash A$  then  $A$  is  $\Gamma\Delta$ -stably true where Y is  $\{k-T\}, \{k-T, k-4\}, \{k-T, k-4, k-5\}$  and Z is  $\{b-4\}, \{b-4, b-5\}$  and  $\Gamma$  is respectively T-, S4-, S5-model and  $\Delta$  is KD4-, KD45-model respectively.*

The third option is to extend the logic by adding stronger axioms for truth, such as T-T and T-4, provided that the rule of necessitation is weakened. Let  $TKB^m[X, Y, Z]$  be the logic  $TKB^m[Y, Z]$  as above where X now is any of  $\{T-T\}, \{T-T, T-4\}$  safe axioms of truth, together with the following rules:

- T-RN. if  $TKB^m[Y, Z] \vdash A$  then  $TKB^m[X, Y, Z] \vdash T(A)$
- k-FN. if  $TKB^m[X, Y, Z] \vdash A$  then  $TKB^m[X, Y, Z] \vdash Know_i(A)$
- b-FN. if  $TKB^m[X, Y, Z] \vdash A$  then  $TKB^m[X, Y, Z] \vdash Bel_i(A)$

**Theorem 13** *If  $TKB^m[X, Y, Z] \vdash A$  then  $A$  is  $\Gamma\Delta$ -stably true where X is any combination of the safe axioms of truth, Y is  $\{k-T\}, \{k-T, k-4\}, \{k-T, k-4, k-5\}$ , Z is  $\{b-4\}, \{b-4, b-5\}$  and  $\Gamma$  is a T-, S4-, S5-model and  $\Delta$  is a KD4-, KD45-model respectively.*

**Theorem 14** *The axioms:*

- kt1.  $Know_i(T(A)) \Leftrightarrow T(Know_i(A))$
- kt2.  $Know_i(A) \Rightarrow Know_i(T(A))$
- bt1.  $Bel_i(T(A)) \Leftrightarrow T(Bel_i(A))$
- bt2.  $Bel_i(A) \Rightarrow Bel_i(T(A))$

*can be consistently added to the family of logics  $TKB^m[X, Y, Z]$ .*

## 7 Syntactic Modalities

The language  $\mathcal{L}_1$  is extended to  $\mathcal{L}_2$  by adding two new predicates  $K_i$  and  $B_i$  for knowledge and belief respectively. These are indexed by  $i$  in order to identify the individual agent. The following two clauses are added to the language:

- (i) if  $t$  is a term then  $K_i(t)$  is a wff
- (ii) if  $t$  is a term then  $B_i(t)$  is a wff

However, such a direct treatment of knowledge and belief as predicates as Montague [39] and Thomason [51] showed, results in inconsistency. These results seem to show that syntactical treatments of the propositional attitudes are not viable. In the presence of self-reference, axiom schemas characterising knowledge and even weaker schemas characterising belief, give rise to inconsistency. As a consequence, none of the standard logics T, S4 and S5 for knowledge and KD4 and KD45 for belief are consistent as logics for the syntactic modalities.

Turner [53] suggests an alternative approach. This involves defining a syntactic modality (predicate) in terms of a classical modal (sentential) operator and the truth predicate. Following this approach and using the language  $\mathcal{L}_1$  two predicates  $K_i$  and  $B_i$  are defined:

$$K_i(t) =_{def} Know_i(T(t))$$

$$B_i(t) =_{def} Bel_i(T(t))$$

Model theoretically these definitions are supported by the following clauses:

$$M \models_{v,w} K_i(A) \text{ iff for all } w' \text{ such that } \mathcal{K}_i(w, w') \mathcal{T}(\| A \|_{v,w'}) = 1$$

$$M \models_{v,w} B_i(A) \text{ iff for all } w' \text{ such that } \mathcal{B}_i(w, w') \mathcal{T}(\| A \|_{v,w'}) = 1$$

Now a translation mechanism between the language of syntactic modality  $\mathcal{L}_2$  and the classical language of modal logic  $\mathcal{L}_1$  can be employed (the subscript  $i$  denoting an agent has been omitted from the modal operators and the syntactic modalities):

- (1)  $trans(x) = x$
- (2)  $trans(\ulcorner A \urcorner) = \ulcorner trans(A) \urcorner$
- (3)  $trans(K(t)) = Know(T(trans(t)))$
- (4)  $trans(B(t)) = Bel(T(trans(t)))$
- (5)  $trans(T(t)) = T(trans(t))$
- (6)  $trans(A \wedge B) = trans(A) \wedge trans(B)$
- (7)  $trans(\neg A) = \neg trans(A)$
- (8)  $trans(\forall x A) = \forall x trans(A)$

Similarly for the other connectives and quantifier.

## 8 Logics for the Syntactic Modalities

The language that will be used to write down the logics of truth and syntactic modality is  $\mathcal{L}_3$ .  $\mathcal{L}_3$  consists of the truth predicate and the syntactic modalities for knowledge and belief. The translation mechanism introduced in the previous section between  $\mathcal{L}_2$  and  $\mathcal{L}_1$  can be employed to determine the consistent logics for the syntactic modalities. From this point onwards, the following naming convention is adopted:

- (i) the names of axioms and rule for the syntactic modality  $K_i$  will be indicated by a capital letter 'K', followed by a dash and the name of the axiom or rule;
- (ii) the names of axioms and rule for the syntactic modality  $B_i$  will be indicated by a capital letter 'B', followed by a dash and the name of the axiom or rule.

The first combined system is called STKB and the strongest axiom of truth, syntactic knowledge and syntactic belief is D:

**Truth**

T-K.  $T(A \Rightarrow B) \Rightarrow (T(A) \Rightarrow T(B))$

T-D.  $T(A) \Rightarrow \neg T(\neg A)$

T-B.  $\forall x T(A) \Rightarrow T(\forall x A)$

T-N. If  $\vdash A$  then  $\vdash T(A)$

**Knowledge**

K-T.  $K_i(A \Rightarrow B) \Rightarrow (K_i(A) \Rightarrow K_i(B))$

K-D.  $K_i(A) \Rightarrow \neg K_i(\neg A)$

K-B.  $\forall x K_i(A) \Rightarrow K_i(\forall x A)$

K-N. If  $\vdash A$  then  $\vdash K_i(A)$

**Belief**

B-K.  $B_i(A \Rightarrow B) \Rightarrow (B_i(A) \Rightarrow B_i(B))$

B-D.  $B_i(A) \Rightarrow \neg B_i(\neg A)$

B-B.  $\forall x B_i(A) \Rightarrow B_i(\forall x A)$

B-N. If  $\vdash A$  then  $\vdash B_i(A)$

**Theorem 15** *STKB is a consistent logic of truth and syntactic modalities.*

As with the logics of truth and classical modality there are three ways in which STKB can be extended: by adding i) stronger axioms for truth; ii) stronger axioms for syntactic knowledge; and iii) stronger axioms for syntactic belief. However, given the definition of the modal predicates in terms of the respective modal sentential operators and the truth predicate, an additional difficulty arises. The axioms for the modal predicates are affected by the same problems as the axioms of truth. Caution is required in order to obtain consistent logics for the syntactic modalities as well. Results regarding logics of truth are known from the previous sections.

Let  $STKB[X, Y, Z]$  be the logic STKB as defined previously and X be a subset of the safe axioms of truth  $\{T-T, T-4\}$ , Y be a subset of the axioms of syntactic knowledge  $\{K-T, K-4\}$  and Z be the axiom of syntactic belief  $\{B-4\}$ . Furthermore, since now the syntactical axiomatisations are infected by the problems of the truth predicate, K-T, K-4 and B-4 are rendered safe, but not stably true. Consequently, the necessitation rules for knowledge and belief need to be weakened:

T-RN. if  $STKB \vdash A$  then  $STKB[X, Y, Z] \vdash T(A)$

K-RN. if  $STKB \vdash A$  then  $STKB[X, Y, Z] \vdash K_i(A)$

B-RN. if  $STKB \vdash A$  then  $STKB[X, Y, Z] \vdash B_i(A)$

**Theorem 16** *The  $STKB[X, Y, Z]$  family of logics is consistent.*

**Theorem 17** (a) *The following axioms linking knowledge with truth can be consistently added to the  $STKB[X, Y, Z]$  logics:*

KT1.  $K_i(T(A)) \Leftrightarrow T(K_i(A))$

KT2.  $K_i(A) \Rightarrow K_i(T(A))$

(b) *The following axioms linking belief with truth can be consistently added to the  $STKB[X, Y, Z]$  logics:*

BT1.  $B_i(T(A)) \Leftrightarrow T(B_i(A))$

*BT2.*  $B_i(A) \Rightarrow B_i(T(A))$

(c) *The axioms KT1 and KT2 allow the derivation of the following:*

$K_i(A) \Leftrightarrow K_i(T(A)) \Leftrightarrow T(K_i(A))$

(d) *The axioms BT1 and BT2 allow the derivation of the following:*

$B_i(A) \Leftrightarrow B_i(T(A)) \Leftrightarrow T(B_i(A))$

The equivalences (c) and (d) seem natural properties for knowledge and belief. If an agent knows something, then it knows it is true and it is true that it knows it; similarly for belief.

This concludes the presentation of the combined systems for predicate truth and the syntactic modalities of knowledge and belief. Additional systems can be uncovered by utilising the non-standard axioms of truth considered by Turner [53], this is left open to investigation.

## 9 Knowledge and Belief

The logics introduced in the previous sections have independent modal or syntactic operators for the epistemic notion of knowledge and the doxastic notion of belief. Axiom schemas that connect knowledge with truth and belief with truth were introduced, but no such axiom schemas relating knowledge with belief were considered. In fact, the accessibility relations that are used in the model for knowledge and belief are completely independent of each other. This separation of knowledge and belief was introduced on purpose so that one need not have to worry during the construction of the framework about their relation.

The nature of knowledge and belief as well as their connection and differences, have been heavily debated in the philosophical literature, without however producing any definitive answers [1, 20, 23, 48]. As used by human agents they are both information-bearing attitudes. However, knowledge of a fact presupposes that the fact is true, whereas a belief is not necessarily always true. The verb “believe” is considered to have a psychological flavour since the state of believing is connected with the internal psychological state of the agent [8]. On the other hand, the verb “know” is not connected with the internal state of an agent, but clearly with the state of the world. A very common definition of knowledge in Philosophy which has been adopted in Computer Science and in Artificial Intelligence as well, is that of knowledge being “true belief” or “justified true belief”. But as Gettier [18] argued, this may not always be the case and truth and belief may be necessary, but not sufficient conditions to ascribe knowledge to an agent. This criterion of knowledge being just true beliefs gives too generous a definition of knowledge; an agent might happen to have some true beliefs that would not be reasonable to characterise as knowledge. Imagine an agent that always believes that it is noon; it would not be reasonable to say that each day at noon the agent has knowledge that it is noon.

Although the aim here is not to clarify the issues on the nature and relationship of knowledge and belief, the following sections will discuss possible connection axiomatisations for the two attitudes. Since computational agents is the main focus of this paper, the verb “believe” does not have a psychological flavour. Instead, whenever an agent believes something, this has an air of “doubt” about it, the agent is not absolutely confident that what it believes, holds true. It might find out later that it was mistaken and in fact its belief was false. On the other hand, knowledge cannot be allowed to be false. Whatever the agent knows to be true must hold true as well. Before proceeding with the presentation of the connection axiomatisations, the following subsection discusses related work on systems of knowledge and belief.

## 9.1 Related Work on Systems of Knowledge and Belief

Kraus and Lehmann’s [31] formalisation is probably one of the first attempts to combine both information attitudes in the same system. The system that they introduce is based on a propositional modal language with two modal operators  $K$  and  $B$ , for knowledge and belief respectively. Additional operators are introduced to express group attitudes. Semantics to the language is given in terms of possible worlds with underlying knowledge and belief accessibility relations  $\mathcal{K}_i$  and  $\mathcal{B}_i$  respectively.  $\mathcal{K}_i$  is an equivalence relation and  $\mathcal{B}_i$  satisfies the conditions:

- i)  $\mathcal{B}_i$  is serial
- ii)  $\mathcal{B}_i$  is contained in  $\mathcal{K}_i$ , that is  $\mathcal{B}_i \subseteq \mathcal{K}_i$
- iii)  $\forall w, w', w'' \in W$ , if  $\mathcal{K}_i(w, w')$  and  $\mathcal{B}_i(w', w'')$ , then  $\mathcal{B}_i(w, w'')$

Kraus & Lehmann’s system (KL) is essentially a combination of an S5 system for knowledge and a KD45 system for belief. There are two interesting axioms relating knowledge and belief:

$K_i\phi \Rightarrow B_i\phi$ , an agent believes everything it knows.

$B_i\phi \Rightarrow K_iB_i\phi$ , an agent knows its own beliefs, i.e. beliefs are conscious.

It is also suggested that an agent could believe that its beliefs amount to knowledge:

(\*)  $B_i\phi \Rightarrow B_iK_i\phi$

However, incorporating such an axiom into KL would result in knowledge collapsing to belief. van der Hoek [54] attempts to find systems where (\*) would be allowed without yielding  $B_i\phi \Rightarrow K_i\phi$ , and reaches the conclusion that the two accessibility relations should not correspond. (\*) is only reasonable if  $B_i\phi$  stands for “ $i$  is convinced that  $\phi$ ” and not for “ $i$  presumes  $\phi$ ”. A disadvantage of the KL system is that  $B_iK_i\phi \Rightarrow K_i\phi$  is a derived theorem which may contradict some of the intuitions about belief, since an agent may believe that it knows  $\phi$  without however  $\phi$  being true.

Voorbraak [58, 59] concentrates on rather idealised notions of knowledge and belief which he calls Objective Knowledge and Rational Belief. Therefore the operator  $K_a\phi$  for knowledge intuitively is read as “the agent  $a$  objectively knows  $\phi$ ”, that is the agent knows  $\phi$  if  $\phi$  is true in every world that is possible according to the information available to  $a$ . The modal operator  $B_a\phi$  means “agent  $a$  is rationally convinced that  $\phi$ ”, that is  $\phi$  is valid in every world that is considered possible by  $a$ . This notion of rational belief is very strong however, and in reality most of the times a much weaker notion of belief is used. Semantics to the propositional language considered by Voorbraak is given not in terms of classical possible worlds but by using generalised functional models [58, 59]. As it is shown the generalised functional model for knowledge corresponds with Kripke models for the modal system S5 while the one for belief corresponds with Kripke models for modal logic KD45. The accessibility relation for knowledge as in the modal logic S5 is an equivalence relation and the accessibility relation for belief is serial, euclidean and transitive as in the KD45 logic. The following are the connection axioms adopted by Voorbraak:

$B(\phi) \Rightarrow K(B(\phi))$

$\neg B(\phi) \Rightarrow K(\neg B(\phi))$

$B(\phi) \Rightarrow B(K(\phi))$

The OK&RIB system differs from the KL system proposed by Kraus and Lehmann [31]. The theorem  $B(\phi) \Rightarrow B(K(\phi))$  that causes knowledge to collapse with belief in KL is present in Voorbraak’s system without causing problems. He argues that the definition of knowledge as “justified true belief” or “true belief” leads to problems and thus the theorem  $K_i\phi \Rightarrow B_i\phi$  is

not valid in OK&RIB. Knowledge and belief are quite separate notions and even if a belief is true it does not always constitute knowledge. However, Voorbraak’s system deals with notions of knowledge and belief that may be considered to be quite strong. Especially the notion of belief is very strong for an agent that is not infallible.

Lammare and Shoham [33] aimed at capturing the connection between the two notions and their relation with certainty. They based their approach upon the idea that knowledge entails certainty which in turn entails belief. Knowledge for the agent is facts for which it is certain, however there is no such connection between certainty and belief. They strongly reject the negative introspection property for a logic of knowledge certainty and belief, since negative introspection seems to result in counterintuitive properties because of the way these notions are defined.

Shoham and Moses [42] depart from the customary view that belief is the primitive notion and knowledge can then be defined in terms of it, and instead take knowledge as primitive and consider belief to be derivable. They accept the S5 modal system for knowledge, but “an agent believes  $\phi$ ” is taken to mean “the agent knows that either  $\phi$ , or else something specific unusual is the case”. In other words, they define belief as knowledge relative to an assumption and the assumption is taken to be defeasible. As part of the definition the assumptions that the agent used to form a particular belief are included, however the choice of assumptions seems to be arbitrary.

Israel [27] accepts only the schema that knowledge requires belief for a logic that combines the epistemic and doxastic notions:

$$K_i(\phi) \Rightarrow B_i(\phi)$$

whereas he rejects:

$$B_i(\phi) \Rightarrow K_i(B_i(\phi))$$

$$K_i(\phi) \Rightarrow B_i(K_i(\phi))$$

In van der Hoek et al. [55], the authors present a rich framework for formalising the properties of reasoning agents. Four modal operators are introduced:  $B_i^k$ ,  $B_i^o$ ,  $B_i^c$ , and  $B_i^d$ , for expressing what the agent knows, believes to be true by observation, believes through communication and believes by default respectively. Knowledge is regarded as information that the agent has been designed with such as domain specific facts, and is considered to be fixed and true. Observational beliefs is information that the agent observes and are regarded to be true. Communicational and default beliefs are not required to be true. The following connect the four attitudes:

$$(B_i^o\phi \Rightarrow \phi) \wedge (B_i^k\phi \Rightarrow \phi)$$

$$(B_i^k\phi \Rightarrow B_i^o\phi) \wedge (B_i^o\chi \Rightarrow B_i^c\chi) \wedge (B_i^c\psi \Rightarrow B_i^d\psi)$$

The knowledge and observational belief operators have the properties of the S5 logic, whereas the communicational and default belief operators those of K45.

## 10 Connection Axiomatisations for Modal Knowledge and Belief

In the following subsections three possible connection axiomatisations for modal knowledge and belief are presented. The first two are considered in the strongest combined system of truth and modal knowledge and belief which is  $\text{TKB}^m[\{\text{T-T}, \text{T-4}\}, \{\text{k-T}, \text{k-4}, \text{k-5}\}, \{\text{b-4}, \text{b-5}\}]$ , while the third in the weaker  $\text{TKB}^m[\{\text{T-T}, \text{T-4}\}, \{\text{k-T}, \text{k-4}\}, \{\text{b-4}, \text{b-5}\}]$ . The results regarding the connection axioms in the modal case are applicable to simple combined modal systems of knowledge and belief. The relevant proofs are provided in Appendix A.1.

## 10.1 Weak Modal Connection

Since knowledge and belief are both information-bearing attitudes, it is reasonable to accept that the weakest form of relation between them is that of consistency:

$$\text{Wkb. } Bel_i(A) \Rightarrow \neg Know_i(\neg A)$$

This axiom expresses that knowledge and beliefs are consistent and hence, if an agent believes  $A$ , then at least it does not know  $\neg A$ . This seems to be a reasonable constraint and in line with our intuitions about knowledge and belief. This axiom schema requires that the intersection of the set of knowledge-accessible and belief-accessible worlds is not the empty set:

**Lemma 18** *Wkb is sound in all models that satisfy the semantic condition:*

$$\text{WSC. } \forall w, \exists w' \quad \mathcal{K}_i(w, w') \wedge \mathcal{B}_i(w, w')$$

Additional connection axioms can be considered in order to refine the relationship between knowledge and belief. For instance, one may accept the principle that if an agent believes  $A$ , then it knows that it believes it, which is a form of introspection:

$$\text{kb1. } Bel_i(A) \Rightarrow Know_i(Bel_i(A))$$

**Lemma 19** *kb1 is sound in all models that satisfy the semantic condition:*

$$\text{C1. } \forall w, w', w'' \quad \mathcal{K}_i(w, w') \wedge \mathcal{B}_i(w', w'') \Rightarrow \mathcal{B}_i(w, w'')$$

Another possible connection axiom for knowledge and belief states that if an agent believes  $A$ , then it believes that it knows  $A$  as well:

$$\text{kb2. } Bel_i(A) \Rightarrow Bel_i(Know_i(A))$$

**Lemma 20** *kb2 is sound in all models that satisfy the semantic condition:*

$$\text{C2. } \forall w, w', w'' \quad \mathcal{B}_i(w, w') \wedge \mathcal{K}_i(w', w'') \Rightarrow \mathcal{B}_i(w, w'')$$

This is another form of introspection, but it is stronger than that expressed by the kb1 axiom since now the agent considers its beliefs to amount to knowledge. Furthermore, one can accept a form of negative introspection regarding beliefs according to which an agent that does not believe  $A$ , believes that it does not also know  $A$ :

$$\text{kb3. } \neg Bel_i(A) \Rightarrow Bel_i(\neg Know_i(A))$$

**Lemma 21** *kb3 is sound in all models that satisfy the semantic condition:*

$$\text{C3. } \forall w, w', w'' \quad \mathcal{B}_i(w, w') \wedge \mathcal{B}_i(w, w'') \Rightarrow \mathcal{K}_i(w', w'')$$

**Proposition 22** *The following formulae are theorems in a modal system of knowledge (S5) and belief (KD45) that satisfies the semantic conditions WSC, C1, C2 and C3.*

1.  $Bel_i(A) \Leftrightarrow Know_i(Bel_i(A))$
2.  $Bel_i(A) \Leftrightarrow Bel_i(Know_i(A))$
3.  $\neg Bel_i(A) \Leftrightarrow Know_i(\neg Bel_i(A))$
4.  $\neg Bel_i(A) \Leftrightarrow Bel_i(\neg Know_i(A))$

## 10.2 Strong s1 Modal Connection

Most of the philosophical approaches presume a strong connection between knowledge and belief: knowledge always implies belief. Such an approach is adopted for instance in Kraus and Lehmann's [31] system. Axiomatically this is captured by the following:

$$\text{Skb. } \text{Know}_i(A) \Rightarrow \text{Bel}_i(A)$$

Semantically this axiom requires that the accessibility relation for belief is contained in the accessibility relation for knowledge,  $\mathcal{B}_i \subset \mathcal{K}_i$ . Thus, the belief-accessible worlds are a subset of the knowledge-accessible worlds.

**Lemma 23** *Skb is sound in all models that satisfy the semantic condition:*

$$\text{SC. } \forall w, w' \mathcal{B}_i(w, w') \Rightarrow \mathcal{K}_i(w, w')$$

The following result is obtained for the logics of modalities and predicate truth:

**Theorem 24** *The following are valid formulae connecting modal knowledge and belief with predicate truth in the  $\text{TKB}^m[X, Y, Z]$  family of logics with the Skb axiom.*

1.  $\text{Know}_i(A) \Rightarrow \text{Bel}_i(T(A))$
2.  $\text{Know}_i(A) \Rightarrow T(\text{Bel}_i(A))$

As in the weak connection axiomatisation, additional properties can be considered in order to refine the relationship between knowledge and belief. For instance, the axiom that expresses that an agent knows all its beliefs can be adopted here as well by imposing the semantic condition C1 (Lemma 19):

$$\text{kb1. } \text{Bel}_i(A) \Rightarrow \text{Know}_i(\text{Bel}_i(A))$$

One may also consider adding the kb2:

$$\text{kb2. } \text{Bel}_i(A) \Rightarrow \text{Bel}_i(\text{Know}_i(A))$$

This requires the semantic condition C2 (Lemma 20) between accessible worlds. However if C2 is imposed, this has the undesired consequence that knowledge and belief collapse into the same attitude ( $\text{Bel}_i(A) \Leftrightarrow \text{Know}_i(A)$ ):

**Theorem 25** ([54]) *If the logic of knowledge and belief satisfies the following conditions:*

$$(i) \forall w, w', w'' \mathcal{B}_i(w, w') \wedge \mathcal{K}_i(w', w'') \Rightarrow \mathcal{B}_i(w, w'')$$

$$(ii) \forall w, w' \mathcal{B}_i(w, w') \Rightarrow \mathcal{K}_i(w, w')$$

$$(iii) \mathcal{B}_i \text{ is serial, i.e. } \forall w \exists w' \mathcal{B}_i(w, w')$$

$$(iv) \mathcal{K}_i \text{ is Euclidean, i.e. } \forall w, w', w'' \mathcal{K}_i(w, w') \wedge \mathcal{K}_i(w, w'') \Rightarrow \mathcal{K}_i(w', w'')$$

*then it is also the case that  $\forall w, w' \mathcal{K}_i(w, w') \Rightarrow \mathcal{B}_i(w, w')$*

Thus, if the kb2 schema is desirable in a logic also containing the strong connection axiom between knowledge and belief, one of the above conditions (iii)-(iv) needs to be abandoned in order to avoid the correspondence of knowledge with belief. van der Hoek [54] refers to this as the *Believed Consciousness of Beliefs* problem. Thus, kb2 will not be considered here as part of the strong s1 connection axiomatisation.

One may consider adding the principle that an agent that does not believe  $A$ , believes that it does not also know  $A$ :

$$\text{kb3. } \neg \text{Bel}_i(A) \Rightarrow \text{Bel}_i(\neg \text{Know}_i(A))$$

Unfortunately, if kb3 is adopted as an axiom then  $\neg \text{Bel}_i(A) \Leftrightarrow \text{Bel}_i(\neg \text{Know}_i(A))$  (\*) is a theorem of the system.  $\neg \text{Know}_i(A) \Rightarrow \text{Bel}_i(\neg \text{Know}_i(A))$  is also a theorem (Proposition 26)



and by modus ponens  $Bel_i(\neg Know_i(A))$ . From this and (\*)  $Bel_i(\neg Know_i(A)) \Rightarrow \neg Bel_i(A)$ , and by modus ponens  $\neg Bel_i(A)$  is obtained. Thus,  $\vdash \neg Know_i(A) \Rightarrow \neg Bel_i(A)$  and by contraposition  $\vdash Bel_i(A) \Rightarrow Know_i(A)$ . As a result, knowledge collapses into belief, and therefore the kb3 axiom is not part of the strong s1 connection axiomatisation for modal knowledge and belief.

**Proposition 26** *The following formulae are theorems in a modal system of knowledge (S5) and belief (KD45) that satisfies the semantic conditions SC and C1.*

1.  $Bel_i(A) \Leftrightarrow Know_i(Bel_i(A))$
2.  $Know_i(A) \Leftrightarrow Bel_i(Know_i(A))$
3.  $\neg Bel_i(A) \Leftrightarrow Know_i(\neg Bel_i(A))$
4.  $\neg Know_i(A) \Leftrightarrow Bel_i(\neg Know_i(A))$
5.  $Know_i(A) \Rightarrow Know_i(Bel_i(A))$
6.  $Know_i(A) \Rightarrow Bel_i(Bel_i(A))$

### 10.3 Strong s2 Modal Connection

The accessibility relation in the previous strong s1 connection axiomatisation between knowledge and belief was taken to be an equivalence relation. As was shown in Theorem 25, if one wishes to incorporate the kb2 axiom between knowledge and belief and certain conditions apply, the notion of knowledge collapses to that of belief. In the s2 connection axiomatisation the accessibility relation for knowledge is taken to be reflexive, serial and transitive, and thus the k-5 axiom for knowledge (negative introspection) is not included. The strong connection axiom is adopted:

$$\text{Skb. } Know_i(A) \Rightarrow Bel_i(A)$$

The result of Theorem 24 pertains here as well since  $TKB^m[\{T-T, T-4\}, \{k-T, k-4\}, \{b-4, b-5\}]$  is one of the logics in the  $TKB^m[X, Y, Z]$  family. Furthermore, the semantic conditions C1 and C2 (Lemmas 19 and 20) are imposed and accordingly the kb1 and kb2 axioms are adopted:

$$\text{kb1. } Bel_i(A) \Rightarrow Know_i(Bel_i(A))$$

$$\text{kb2. } Bel_i(A) \Rightarrow Bel_i(Know_i(A))$$

Since the negative introspection axiom for knowledge is rejected, two theorems of the s1 connection can only be introduced as axioms:

$$\text{kb4. } \neg Know_i(A) \Rightarrow Bel_i(\neg Know_i(A))$$

$$\text{kb5. } \neg Bel_i(A) \Rightarrow Know_i(\neg Bel_i(A))$$

**Lemma 27** *a) kb4 is sound in all models that satisfy the semantic condition:*

$$C4. \quad \forall w, w', w'' \quad \mathcal{K}_i(w, w') \wedge \mathcal{B}_i(w, w'') \Rightarrow \mathcal{K}_i(w', w'')$$

*b) kb5 is sound in all models that satisfy the semantic condition:*

$$C5. \quad \forall w, w', w'' \quad \mathcal{B}_i(w, w') \wedge \mathcal{K}_i(w, w'') \Rightarrow \mathcal{B}_i(w', w'')$$

The kb3 axiom ( $\neg Bel_i(A) \Rightarrow Bel_i(\neg Know_i(A))$ ), which in the strong s1 connection axiomatisation causes knowledge to collapse into belief, cannot be included here.

**Proposition 28** *The following formulae are theorems in a modal system of knowledge (S4) and belief (KD45) that satisfies the semantic conditions SC, C1, C2, C4 and C5.*

1.  $Bel_i(A) \Leftrightarrow Know_i(Bel_i(A))$

2.  $Bel_i(A) \Leftrightarrow Bel_i(Know_i(A))$
3.  $\neg Bel_i(A) \Leftrightarrow Know_i(\neg Bel_i(A))$
4.  $\neg Know_i(A) \Leftrightarrow Bel_i(\neg Know_i(A))$
5.  $Know_i(A) \Rightarrow Know_i(Bel_i(A))$
6.  $Know_i(A) \Rightarrow Bel_i(Bel_i(A))$

The three connection axiomatisations for modal knowledge and belief were investigated along the lines of the strongest systems for knowledge and belief that do not allow the two notions to collapse. Intermediate results can be obtained for weaker systems of knowledge and belief. Space does not allow the presentation of all these weaker systems, but the reader can get a flavour of the systems available in Appendix A.2.

## 11 Connection Axiomatisations for Syntactic Knowledge and Belief

Based upon the connection axiomatisations for modal knowledge and belief the following sections present connection axiomatisations for the respective syntactic modalities. Notably, the rules of necessitation for both syntactic modalities are weakened in order to retain consistency. The axiomatisations are considered in the context of the strongest system  $STKB[\{T-T, T-4\}, \{K-T, K-4\}, \{B-4\}]$ .

### 11.1 Weak Syntactic Connection

Following the same route as in the modal case, a weak connection axiomatisation between syntactic knowledge and belief is first considered.

**Proposition 29** *The following axioms can be consistently added to  $STKB[\{T-T, T-4\}, \{K-T, K-4\}, \{B-4\}]$  in order to define a weak connection.*

1. *WKB.*  $B_i(A) \Rightarrow \neg K_i(\neg A)$
2. *KB1.*  $B_i(A) \Rightarrow K_i(B_i(A))$
3. *KB2.*  $B_i(A) \Rightarrow B_i(K_i(A))$
4. *KB3.*  $\neg B_i(A) \Rightarrow B_i(\neg K_i(A))$

The resulting system is prefixed by ‘W-’ to indicate that it includes the weak syntactic connection axioms.

**Proposition 30** *The following are theorems in  $W-STKB[\{T-T, T-4\}, \{K-T, K-4\}, \{B-4\}]$ :*

1.  $B_i(A) \Leftrightarrow K_i(B_i(A))$
2.  $B_i(\neg K_i(A)) \Rightarrow \neg K_i(A)$

In contrast to the modal case, the following are non-theorems:

- $\not\vdash \neg B(A) \Rightarrow K(\neg B(A))$
- $\not\vdash B(K(A)) \Rightarrow B(A)$
- $\not\vdash B(\neg K(A)) \Rightarrow \neg B(A)$

### 11.2 Strong S1 Syntactic Connection

The strong S1 axiomatisation is again considered in the context of the strongest system:

**Proposition 31** *The following axioms can be consistently added to  $STKB[\{T-T, T-4\}, \{K-T, K-4\}, \{B-4\}]$  in order to define a Strong S1 connection between knowledge and belief:*

1. *SKB.*  $K_i(A) \Rightarrow B_i(A)$
2. *KB1.*  $B_i(A) \Rightarrow K_i(B_i(A))$

The resulting system is indicated by the prefix ‘S1-’.

**Proposition 32** *The following are theorems in S1-STKB[{T-T, T-4}], {K-T, K-4}, {B-4}]:*

1.  $B_i(A) \Leftrightarrow K_i(B_i(A))$
2.  $K_i(A) \Rightarrow B_i(K_i(A))$
3.  $K_i(A) \Rightarrow K_i(B_i(A))$
4.  $K_i(A) \Rightarrow B_i(B_i(A))$

Non-theorems include:

- $\not\vdash \neg B_i(A) \Rightarrow K_i(\neg B_i(A))$
- $\not\vdash B_i(K_i(A)) \Rightarrow K_i(A)$
- $\not\vdash \neg K_i(A) \Rightarrow B_i(\neg K_i(A))$

In addition, the following formulae connect knowledge with belief and truth:

**Proposition 33** *The following are theorems in S1-STKB[{T-T, T-4}], {K-T, K-4}, {B-4}]:*

1.  $K_i(A) \Rightarrow B_i(T(A))$
2.  $K_i(A) \Rightarrow T(B_i(A))$

### 11.3 Strong S2 Syntactic Connection

The equivalent of the s2 connection axiomatisation for the syntactic modalities can be obtained as follows:

**Proposition 34** *The following axioms can be consistently added to STKB[{T-T, T-4}], {K-T, K-4}, {B-4}] in order to define a Strong S1 connection between knowledge and belief:*

1. SKB.  $K_i(A) \Rightarrow B_i(A)$
2. KB1.  $B_i(A) \Rightarrow K_i(B_i(A))$
3. KB2.  $B_i(A) \Rightarrow B_i(K_i(A))$
4. KB4.  $\neg K_i(A) \Rightarrow B_i(\neg K_i(A))$

The resulting system is prefixed by ‘S2-’ to indicate that it includes the strong S2 connection axioms.

**Proposition 35** *The following are theorems in S2-STKB[{T-T, T-4}], {K-T, K-4}, {B-4}]:*

1.  $B_i(A) \Leftrightarrow K_i(B_i(A))$
2.  $K_i(A) \Rightarrow B_i(K_i(A))$
3.  $K_i(A) \Rightarrow K_i(B_i(A))$
4.  $K_i(A) \Rightarrow B_i(B_i(A))$
5.  $\neg K_i(A) \Leftrightarrow B_i(\neg K_i(A))$

The formula  $B(K(A)) \Rightarrow B(A)$  is not a theorem and furthermore, the axiom  $\neg B_i(A) \Rightarrow K_i(\neg B_i(A))$  which was imposed on the respective s2 type of connection for modal knowledge and belief, cannot be adopted here since it would render  $\neg B_i(A) \Rightarrow B_i(\neg B_i(A))$  a theorem of the system. (Proof. Assume  $\neg B_i(A)$ , then from the axiom  $\neg B_i(A) \Rightarrow K_i(\neg B_i(A))$  and modus ponens  $K_i(\neg B_i(A))$ . From the connection axiom SKB it follows that  $K_i(\neg B_i(A)) \Rightarrow B_i(\neg B_i(A))$  and by modus ponens  $B_i(\neg B_i(A))$ ) The results of Proposition 33 follow for S2-STKB[{T-T, T-4}], {K-T, K-4}, {B-4}] as well.

As in the modal case, the connection axiomatisations were investigated along the lines of the strongest system for knowledge and belief that does not allow the two notions to collapse. Intermediate results can be obtained for weaker systems of knowledge and belief and the reader is referred for more details to Appendix A.2.

## 12 Further Work

The syntactical framework that has been described in this paper offers another possible and attractive way of formalising the properties of agents and their reasoning.

Although the approach followed involves two independent modal (and syntactic) operators for knowledge and belief, the alternative view of knowledge being simply true belief can be accommodated as well. One could start with introducing only one modal operator, that of belief  $Bel_i$  and define a belief syntactic operator  $B_i$  as well as knowledge as follows:

$$B_i(A) =_{def} Bel_i(T(A))$$

$$K_i(A) =_{def} B_i(A) \wedge T(A)$$

As a result, knowledge always implies belief:

$$K_i(A) \Rightarrow B_i(A)$$

As the reader can check the logic for  $K_i$  based on an underlying system S4 for truth and KD4 for belief, is S4 (the necessitation rules are still restricted, as in the case of independent syntactic modalities). By defining knowledge as true belief though, it is not possible to adopt a weak connection axiomatisation as when independent operators are used.

Apart from knowledge and belief the framework can be used to formalise other propositional attitudes such as desires and intentions. In particular, the use of a syntactical approach may address some of the issues regarding intentions as discussed in [30]. Some initial investigations into formalising desires and intentions have been carried out in [14, 15]. Instead of BDI agents one may want to consider KBDI agents, agents that are characterised by knowledge, beliefs, desires and intentions. The obvious question that springs to mind is why choose to represent agents as having both knowledge and beliefs. Obviously, if one's aim is to formalise theories of artificial agents, it is natural to draw one's inspiration from human agents. Even if human agents may not turn out to be very good models in all situations, nevertheless they provide some guiding principles. On the other hand, highly advanced artificial agents are expected to approximate human commonsense reasoning, and therefore almost necessarily they must be ascribed a variety of human attitudes. In this endeavour information is a key issue, since it represents what the agent perceives about the world which will then use to change this world. The use of both knowledge and belief may provide a better understanding of agents which may lead to more concrete specifications and even more realistic or "believable" models of agents. The attribution of knowledge, beliefs, desires and intentions leads to conceptually richer frameworks capable of characterising agents with diverse features. Among the most interesting issues open up to investigation are:

i) What is the connection between knowledge-belief and intentions? Surely intentions are quite different from desires since they presuppose that the agent is committed towards their fulfillment. What is the role of knowledge in an agent's decision making process and how is that different to that of belief?

ii) What is the connection between knowledge-belief and the desires of an agent? Would it be possible that the agent's desires are completely decoupled from its knowledge and only related to its beliefs?

In particular, providing notions of realism [47] for KBDI agents, that is a set of connection axioms relating all four notions, is a non-trivial problem. There may not be a unique notion of realism that describes the ideal KBDI agent, since there may be no definitive answer to what an ideal KBDI agent is.

Another attitude that could be incorporated is that of obligation. As with desires and

intentions, this cannot be done willy-nilly, since the relation of the additional attitude with the existing ones needs to be sorted out at the modal level. The notion of obligation poses some interesting questions in relation to the other attitudes. For instance, it may be reasonable to require an agent that is obligated to bring about a state of affairs to intend it as well, or at least to require that the agent should not intend its negation.

Moreover, an agent living in a multi-agent world needs to be able to express and reason about social attitudes such as common knowledge and mutual belief as well a joint intentions and social commitments. Another useful extension would be to enrich the framework with such attitudes.

In the discussion so far the world was regarded as static, no temporal representation has been considered. But the world is inherently dynamic and this needs to be reflected on the formalism. Despite its shortcomings, situation calculus [38] is probably the simplest way of incorporating the notion of time into this framework.

The theory would also benefit from a representation of actions which are the means by which agents attempt to achieve their objectives and change the world around them. Actions are entities of some sort, and entities are usually formally represented by singular terms. Despite actions being regarded as objects, they have an internal structure that other objects do not. One would like to be able to quantify over actions as it is possible with other ordinary entities. Action statements need to be extensional in order to conform to the principle of substitution of equals. For instance, if agent *Bill* kills *John*, and *John* is *Bill's* neighbour, then surely *Bill* has killed his neighbour is true as well. Nevertheless, there may be times when actions need to be treated as intensional entities. That is to say that the way an action is represented may be important: in general knowing how to perform an action does not imply knowing how to perform the same action if it is described differently. One possible way of representing actions is to adopt the approach advocated by Davies [7]. Davies considers actions as individuals which are represented as terms which denote a restricted class of sentences of the logical language. This would accord well with the rest of the theory here, since the mechanism for encoding sentences as terms in the language is already in place. Complex actions like sequential, conditional, actions performed in while loops, weak and strong concurrent actions would be introduced as functions. Finally, representing and reasoning about plans is another important extension that would have to be considered.

### 13 Conclusions

This paper concentrated on two main issues: a syntactical treatment for the representation of the propositional attitudes and a comprehensive study of the properties of modal and syntactical systems of knowledge and belief with predicate truth. This work is based upon and extends the work of Turner [53], whereby although propositions are denoted by sentences, the syntactic category of propositions is a subset of the wff of the language. The framework allows for quantification over the propositional attitudes as well as the expression of self-referential statements in the language. To preserve consistency certain measures had to be taken. Firstly, the rules of necessitation for the logics of predicate truth and the syntactic modalities of knowledge and belief had to be restricted to derivations not including the stronger axioms T and S4. Secondly, the negative introspection axiom for both knowledge and belief was discarded from the syntactical systems since it is not safe. Despite these restrictions, the logics are still powerful enough to serve as specifications for agents. Thus, agents can reason about their own knowledge and belief, those of the other agents as well as

what holds true.

Knowledge and belief were formalised independently, first as modal operators and then as syntactic modalities. Three connection axiomatisations were considered in the strongest combined systems for modal knowledge and belief that do not allow the two notions to collapse:

i) Weak (w): Knowledge does not necessarily imply belief, but beliefs are consistent with knowledge.

ii) Strong (s1): Knowledge always implies belief. The principle that an agent that believes  $A$ , also believes that it knows  $A$  is not part of this axiomatisation since it causes knowledge and belief to collapse.

ii) Strong (s2): Knowledge always implies belief, but the negative introspection axiom for knowledge is discarded here. Instead, the principle that an agent that believes  $A$ , also believes that it knows  $A$  is accepted.

The inspiration for the three connection axiomatisations between knowledge and belief came mainly from three sources: Kraus and Lehmann [31], van der Hoek [54] and Voorbraak [58]. The weak connection axiomatisation is based upon Voorbraak's ideas for a logic of objective knowledge and rational belief. However, the weak connection axiom that guarantees that whenever an agent believes a proposition at least it should not know its negation ( $B(A) \Rightarrow \neg K(\neg A)$ ), is not part of Voorbraak's axiomatisation. Indeed, the author is not aware of an approach in AI that considers this as the basic connection axiom between knowledge and belief. The strong s1 connection axiomatisation is based on Kraus and Lehmann system. As a consequence, the principle expressing that an agent that believes that it knows  $A$  also knows  $A$  ( $Bel_i(Know_i(A)) \Rightarrow Know_i(A)$ ), is a theorem in the system. This can be alleviated by discarding the negative introspection axiom for knowledge. The s2 strong connection axiomatisation for modal knowledge and belief is based on the observations of van der Hoek regarding the drawbacks of Kraus and Lehmann's system. Since space does not allow the presentation of the properties of weaker systems in combination with the connection axiomatisations, the interested reader is referred to the appendix.

Based on the results for the modal case, connection axiomatisations were considered in the context of logics for the syntactic modalities. Even though syntactical approaches for knowledge and belief have been studied in the literature, none of these involves both propositional attitudes as independent predicates. As a result, connection axiomatisations for syntactic knowledge and belief have not been adequately studied. Such an endeavour presents two main difficulties: preserving consistency and avoiding the collapse of knowledge and belief. As in the modal case three connection axiomatisations were considered:

i) Weak (W): Knowledge does not necessarily imply belief, but beliefs are consistent with knowledge.

ii) Strong (S1): Knowledge always implies belief. The principle that an agent that believes  $A$ , also believes it knows  $A$  is not part of this axiomatisation.

ii) Strong (S2): Knowledge always implies belief, and the principle that an agent that believes  $A$ , also believes it knows  $A$  is included.

Again the reader is referred to [16] for details of weaker syntactical systems in combination with the connection axiomatisations.

The work presented here was motivated partly by the lack of frameworks with sufficient expressive power to allow for quantification over propositions and self-referential statements, and partly by the need to formalise both information attitudes of knowledge and belief as they

express different aspects of an agent’s information about the world. As was argued, agent theories need to be sufficiently expressive to facilitate quantification over the objects of belief, knowledge and other propositional attitudes. Moreover, the concepts of knowledge, truth and belief seem essential for a comprehensive theory of reasoning agents. As a consequence the framework developed here is rather different from those found in the literature. A syntactical approach was advocated which facilitates among other things quantification over propositions as well as the expression of self-referential statements in an elegant way. This allows for a richer set of statements regarding an agent’s reasoning about itself and other agents to be captured. In particular, the example statements of Section 3 which cannot be formalised in propositional or first order modal frameworks such as [5, 55, 47, 60], can be easily formalised in the developed framework. In fact, the formalisation of the example statements of Section 3 was done within this particular framework. Despite the inherent difficulty of retaining consistency while striving for sufficiently strong axiomatisations, the theory yields expressive logics of truth, knowledge and belief. It has also been shown how the interaction of knowledge and belief can be captured while remaining distinct as concepts. This was done both at the modal as well as at the syntactical level. In particular, this investigation shed new light on the interaction of knowledge and belief at the syntactical level; this required care in order to preserve consistency on the one hand and avoid the collapse of knowledge and belief on the other. As far as the author is aware this has not been considered before.

In conclusion the work presented here offers an alternative framework for formalising the properties of agents which is more expressive than other approaches in the literature [5, 55, 47, 60] with respect to quantification over the propositional attitudes and self-referential statements. The framework has the potential to be extended to a comprehensive theory of reasoning agents and although a lot of work remains to be done, this exposition lays the foundations and offers hope in the use of syntactical systems.

## References

- [1] D.M. Armstrong. *Belief, Truth and Knowledge*. Cambridge University Press, 1973.
- [2] N. Asher and H. Kamp. Self-reference, attitudes and paradox. In Partee B.H. Chierchia G. and Turner R., editors, *Properties, Types and Meaning*, volume I, pages 85–158. Kluwer Academic publishers, 1989.
- [3] J. Barwise and J. Etchemendy. *The Liar: An Essay on Truth and Circularity*. Oxford University Press, 1987.
- [4] G. Bealer. *Quality and Concept*. Clarendon Press, 1982.
- [5] P.R. Cohen and H.J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.
- [6] N. Davies. A first order logic of truth, knowledge and belief. In *Proceedings Logics in Artificial Intelligence, European workshop JELIA '90*, pages 170–179, 1990.
- [7] N. Davies. *Truth Modality and Action*. PhD thesis, University of Essex, 1993.
- [8] E. Davis. *Representations of Commonsense Knowledge*. Morgan Kaufmann publishers, 1990.
- [9] D.C. Dennett. *The Intentional Stance*. The MIT Press, 1987.
- [10] J. des Rivieres and H.J. Levesque. The consistency of syntactical treatments of knowledge. *Computational Intelligence*, 4:31–41, 1986.
- [11] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, MA., 1995.

- [12] R. Fagin, J.Y. Halpern, and M.Y. Vardi. Belief, awareness and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.
- [13] R. Fagin, J.Y. Halpern, and M.Y. Vardi. Nonstandard approach to the logical omniscience problem. *Artificial Intelligence*, 79:203–240, 1995.
- [14] M. Fasli. Modeling reasoning agents. In *Proceedings of the 4th Workshop on Practical Reasoning and Rationality (IJCAI'99 Workshop KRR-1)*, pages 8–15, 1999.
- [15] M. Fasli. Modelling agents as reflective intentional systems. In *Proceedings of the First Asia-Pacific Conference on Agent Technology (IAT'99)*, pages 207–211, 1999.
- [16] M. Fasli. Reasoning about knowledge and belief: A syntactical treatment. Technical Report Technical Report CSM-376, University of Essex, Department of Computer Science, 2002.
- [17] S. Feferman. Toward useful type-free theories. *The Journal of Symbolic Logic*, 49(1):75–111, 1984.
- [18] E. Gettier. Is true belief knowledge? *Analysis*, 23:121–123, 1963.
- [19] P.C. Gilmore. The consistency of partial set theory without extensionality. In *Axiomatic Set Theory, Proceedings of Symposia Pure Maths*, volume XIII, pages 147–153. Amer.Math.Soc., 1974.
- [20] A.P. Griffiths. *Knowledge and Belief*. Oxford University Press, 1967.
- [21] H. Gupta. Truth and paradox. *Journal of Philosophical Logic*, 11:1–60, 1980.
- [22] J. Halpern and Y. Moses. Towards a theory of knowledge and ignorance: Preliminary report. In K.R. Apt, editor, *Logics and Models of Concurrent Systems*, pages 459–476. Springer-Verlag, 1985.
- [23] E. Hankins Wolgast. *Paradoxes of Knowledge*. Cornell University Press, 1977.
- [24] H. Herzberger. Notes on naive semantics. *Journal of Philosophical Logic*, 11:61–102, 1982.
- [25] J. Hintikka. *Knowledge and Belief*. Cornell University Press, 1962.
- [26] G.E. Hughes and M.J. Cresswell. *An Introduction to Modal Logic*. Methew & Co Ltd, 1968.
- [27] D. Israel. A weak logic of knowledge and belief: Epistemic and doxastic logic for the yuppie generation. Technical Report Technical Note 359, SRI International, Artificial Intelligence Center, Computer Science and Technology Division, Menlo Park, California, 1985.
- [28] K. Konolige. *A First-Order Formalisation of Knowledge and Action for a Multi-Agent Planning System*, pages 41–72. Number 10. 1982.
- [29] K. Konolige. *A Deduction Model of Belief*. Morgan and Kaufmann Publishers, 1986.
- [30] K. Konolige and M. Pollack. A representationalist theory of intention. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93)*, pages 390–395, 1993.
- [31] S. Kraus and D. Lehmann. Knowledge, belief and time. *Theoretical Computer Science*, 58:155–174, 1988.
- [32] S.A. Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–716, 1975.
- [33] P. Lamarre and Y. Shoham. Knowledge, certainty, belief and conditionalisation. In *Principles of Knowledge Representation and Reasoning, Proceedings of the Fourth International Conference (KR'94)*, pages 415–424, 1994.
- [34] H.J. Levesque. A logic of implicit and explicit belief. In *Proceedings of the National Conference on Artificial Intelligence (AAAI 84)*, pages 198–202, 1984.



- [35] G. McArthur. Reasoning about knowledge and belief: A survey. *Computational Intelligence*, 4:223–243, 1988.
- [36] J. McCarthy. Ascribing mental qualities to machines. In M. Ringle, editor, *Philosophical Perspectives in Artificial Intelligence*, pages 161–195. The Harvester Press Limited, 1979.
- [37] J. McCarthy. First order theories of individual concepts and propositions. In J.E Hayes, D. Michie, and L.I. Miculich, editors, *Machine Intelligence 9*, pages 120–147. Halstead Press, 1979.
- [38] J. McCarthy and P. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence 4*, pages 463–502. 1969.
- [39] R. Montague. The proper treatment of quantification in ordinary english. In Thomason R., editor, *Formal Philosophy, Selected Papers of Richard Montague*, pages 247–170. Yale University Press, 1973.
- [40] R. Montague and D. Kaplan. A paradox regained. *Notre Dame Journal of Symbolic Logic*, 1:79–90, 1960.
- [41] M. Morreau and S. Kraus. Syntactical treatments of propositional attitudes. *Artificial Intelligence*, 106:161–177, 1998.
- [42] Y. Moses and Y. Shoham. Belief as defeasible knowledge. *Artificial Intelligence*, 64:299–321, 1993.
- [43] R. Parikh. Recent issues in reasoning about knowledge. In Parikh R., editor, *Proceedings of Theoretical Aspects of Reasoning About Knowledge Conference (TARK-90)*, pages 3–10, 1990.
- [44] D. Perlis. Languages with self-reference i: Foundations. *Artificial Intelligence*, 25:301–322, 1985.
- [45] D. Perlis. Languages with self-reference ii: Knowledge, belief and modality. *Artificial Intelligence*, 34:179–212, 1988.
- [46] V. Rantala. Impossible worlds semantics and logical omniscience. *Acta Philosophica Fennica*, 35:106–115, 1982.
- [47] A. Rao and M. Georgeff. Decision procedures for bdi logics. *journal of Logic and Computation*, 8(3):293–343, 1998.
- [48] N. Salmon and S. Soames. *Propositions and Attitudes*. Oxford University Press, 1988.
- [49] C. Smorynski. *Self-Reference and Modal Logic*. Springer-Verlag, 1985.
- [50] A. Tarski. *The concept of truth in formalised languages*. Oxford Clarendon Press, 1976.
- [51] R. Thomason. A note on syntactical treatments of modality. *Synthese*, 44:391–395, 1980.
- [52] R. Turner. A theory of properties. *Journal of Symbolic Logic*, 52(2):445–472, 1987.
- [53] R. Turner. *Truth and Modality for Knowledge Representation*. The MIT Press, 1990.
- [54] W. van der Hoek. Systems for knowledge and beliefs. In *Proceedings of the European Workshop in Logics in Artificial Intelligence (JELIA '90)*, volume LNAI:478, pages 267–281. Springer-Verlag, 1990.
- [55] W. van der Hoek, B. van Linder, and J.-J.Ch. Meyer. An integrated modal approach to rational agents. In M. Wooldridge and A. Rao, editors, *Foundations of Rational Agency*, Applied Logic Series 14, pages 133–168. Kluwer, 1998.
- [56] B. van Linder, W. van der Hoek, and J.-J.Ch. Meyer. Formalising motivational attitudes of agents: On preference, goals and commitments. In M. Wooldridge, J.P. Muller, and M. Tambe, editors, *Intelligent Agents II - Agent Theories, Architectures and Languages*, volume LNAI: 1037, pages 17–32. Springer-Verlag, 1996.
- [57] B. van Linder, W. van der Hoek, and J.-J.Ch. Meyer. Formalizing abilities and oppor-

- tunities of agents. *Fundameta Informaticae*, 34:53–101, 1998.
- [58] F. Voorbraak. The logic of objective knowledge and rational belief. In *Proceedings of the European Workshop in Logics in Artificial Intelligence (JELIA '90)*, volume LNAI:478, pages 499–515. Springer-Verlag, 1990.
- [59] F. Voorbraak. Generalized kripke models for epistemic logic. In *Proceedings of the Fourth Conference on Theoretical Aspects of Reasoning about Knowledge (TARK-92)*, pages 214–228, 1992.
- [60] M. Wooldridge. *Reasoning about Rational Agents*. The MIT Press, 2000.

## A Appendix

### A.1 Proofs

**Proof. Theorem 4.** From the Diagonalisation Lemma [49] it is known that there is a wff  $A$  such that:  $A \Leftrightarrow T(\neg A)$ . Then from the necessitation rule  $T(A \Leftrightarrow T(\neg A))$  (1) is obtained. From the tautology  $(A \Rightarrow \neg A) \Rightarrow \neg A$  and necessitation I obtain  $T((A \Rightarrow \neg A) \Rightarrow \neg A)$  (2). From the tautology  $[(A \Rightarrow \neg A) \Rightarrow \neg A] \Rightarrow [(A \Leftrightarrow T(\neg A)) \Rightarrow ((T(\neg A) \Rightarrow \neg A) \Rightarrow \neg A)]$  necessitation yields  $T([(A \Rightarrow \neg A) \Rightarrow \neg A] \Rightarrow [(A \Leftrightarrow T(\neg A)) \Rightarrow ((T(\neg A) \Rightarrow \neg A) \Rightarrow \neg A)])$  (3). From (1),(2), (3) and by the T-K axiom  $T((T(\neg A) \Rightarrow \neg A) \Rightarrow \neg A)$  (4) is obtained. Using the T-T axiom, necessitation and the T-K axiom from (4) I obtain  $T(\neg A)$  (5). By (5) and the T-T axiom  $\neg A$  (6). On the other hand, by (1) and the T-T axiom:  $A \Leftrightarrow T(\neg A)$  (7), which by (5) gives  $A$  (8). But (6) and (8) are contradictory. ■

**Proof. Theorem 5.** Notice that any wff provable in FOPC will be true in all models and therefore stably true. Hence, if  $\vdash_{FOPC} A$  then  $\vdash T(A)$ . ■

**Proof. Theorem 6.** Notice that any wff provable in the D logic will be true in all models and hence stably true. ■

**Proof. Theorem 7.** From Turner [53] it is known that T-T and T-4 are safe and thus they preserve safety. Moreover from Theorem 3 ([53]) and Theorem 6 it is known that the rest of the axioms and inference rule preserve safety as well and furthermore they are stably true. ■

**Proof. Theorem 10.** I use induction on the proofs in  $TKB^m$ . The proof for the axioms of truth follow the same pattern as in the non-modal case. For T-N (of Truth) it is sufficient to observe that if  $A$  is true at every world in every model from some ordinal onwards, then  $T(A)$  will be as well. The axioms for  $Know_i$  are  $\Gamma\Delta$ -stably true since they are true at every world for every  $\Gamma\Delta$ -model. For k-N of knowledge I observe that if  $A$  is true at every world at every model from some ordinal onwards the  $Know_i(A)$  will be. The axioms for the modal operator  $Bel_i$  are  $\Gamma\Delta$ -stably true since they are true at every world for every  $\Gamma\Delta$ -model. For the necessitation rule again b-N if  $A$  is true at every world from some ordinal onwards then  $Bel_i(A)$  will be as well. ■

**Proof. Theorem 11.** For both (a) and (b) It is sufficient to observe that the axioms will be true at all worlds in all models of the appropriate kind. ■

**Proof. Theorem 12.** From the stable truth of the  $TKB^m$  axioms it follows that the axioms are stably true. For the three rules of necessitation it suffices to notice that if  $A$  is true in all worlds from some ordinal onwards, then  $Know_i(A)$ ,  $Bel_i(A)$  and  $T(A)$  will be true as well. ■

**Proof. Theorem 13.** If  $A$  is an instance of any of the axioms of  $TKB^m[Y,Z]$  I employ Theorem 12. Consider the k-FN rule. If  $A$  is true at every world for every stabilization

ordinal then  $Know_i(A)$  will be true as well. The rule for belief follows in a similar way and the rule T-RN for truth follows from the fact that all the theorems of  $TKB^m[Y,Z]$  are  $\Gamma\Delta$ -stable where Y and Z are any combination of the epistemic and doxastic axioms respectively and  $\Gamma$  and  $\Delta$  are the appropriate models. ■

**Proof. Theorem 14.** For kt1, let  $\sigma$  be any stabilisation ordinal. The left hand side of the equivalence amounts to:

$$\forall w \forall w' [\mathcal{K}_i(w, w') \Rightarrow (\forall \beta \geq \sigma) M(\beta) \models_{v, w'} A]$$

and the right-hand side amounts to

$$\forall w \forall \beta \geq \sigma \forall w' [\mathcal{K}_i(w, w') \Rightarrow M(\beta) \models_{v, w'} A]$$

which are obviously equivalent.

For kt2,  $Know_i(A)$  model theoretically is supported by the following clause where  $\alpha, \beta$  are ordinals and  $M_\alpha$  is a stabilised model, i.e.  $\alpha$  is a stabilization ordinal:

$$\forall \beta \geq \alpha \forall w, w' \text{ such that } \mathcal{K}_i(w, w') \text{ then } M_\beta \models_{v, w'} A$$

Then  $Know_i(T(A))$  will similarly be

$$\forall \beta \geq \alpha + 1 \forall w, w' \text{ such that } \mathcal{K}_i(w, w') \text{ then } M_\beta \models_{v, w'} A$$

Notice that in all models in which  $Know_i(A)$  holds,  $Know_i(T(A))$  will hold as well and consequently the axiom can be added to any logic without introducing inconsistency. Similar arguments hold for bt1 and bt2. ■

**Proof. Theorem 15.** The translation function is employed. I illustrate with the K-D axiom and the rule K-N. For the K-D axiom assume  $K_i(A)$ , then by the translation  $Know_i(T(A))$ . By the D-T axiom I obtain  $T(A) \Rightarrow \neg T(\neg A)$  and from necessitation for  $Know$  k-N  $Know_i(T(A) \Rightarrow \neg T(\neg A))$ . By the k-K axiom  $Know_i(T(A)) \Rightarrow Know_i(\neg T(\neg A))$  and now by modus ponens  $Know_i(\neg T(\neg A))$ . From the k-D axiom  $Know_i(\neg T(\neg A)) \Rightarrow \neg Know_i(\neg \neg T(\neg A))$  is obtained, and by double negation  $Know_i(\neg T(\neg A)) \Rightarrow \neg Know_i(T(\neg A))$ . By modus ponens I obtain  $\neg Know_i(T(\neg A))$  and by employing the translation  $\neg K_i(\neg A)$ . For the K-N rule assume  $A$ , and thus by the translation  $trans(A)$ . Then by T-N  $T(trans(A))$  and by k-N  $Know_i(T(trans(A)))$ , and now this translates back to  $K_i(A)$ . ■

**Proof. Theorem 16.** To prove consistency the translation and theorem 15 are employed. I only have to prove consistency of the axioms for the syntactic modalities. For K-T assume  $K_i(A)$  and thus by the translation  $Know_i(T(A))$ . From the k-T axiom  $Know_i(T(A)) \Rightarrow T(A)$  and by modus ponens  $T(A)$ . Now from the T-T axiom  $T(A) \Rightarrow A$  and then by modus ponens  $A$ . For K-4 assume  $K_i(A)$  and thus by the translation  $Know_i(T(A))$ . From the T-4 axiom  $T(A) \Rightarrow T(T(A))$  and by k-N  $Know_i(T(A) \Rightarrow T(T(A)))$ . By k-K I obtain  $Know_i(T(A)) \Rightarrow Know_i(T(T(A)))$  and by modus ponens  $Know_i(T(T(A)))$ . By the k-4 axiom  $Know_i(T(T(A))) \Rightarrow Know_i(Know_i(T(T(A))))$  (i) is obtained. By the kt1 axiom  $Know_i(T(T(A))) \Rightarrow T(Know_i(T(A)))$  and by applying k-N  $Know_i(Know_i(T(T(A))) \Rightarrow T(Know_i(T(A))))$  and now by distribution of knowledge (k-K)  $Know_i(Know_i(T(T(A)))) \Rightarrow Know_i(T(Know_i(T(A))))$  (ii). From (i) and (ii) and modus ponens  $Know_i(T(Know_i(T(A))))$  is obtained and by the translation  $K_i(K_i(A))$ . A similar proof establishes the consistency of B-4. ■

**Proof. Theorem 17.** (a) The translation is employed. For KT1 it yields  $Know_i(T(T(A))) \Leftrightarrow T(Know_i(T(A)))$  which is an instance of the sentential axiom  $Know_i(T(A)) \Leftrightarrow T(Know_i(A))$ . For KT2 by employing the translation I get  $Know_i(T(A)) \Rightarrow Know_i(T(T(A)))$  which is just an instance of the sentential axiom kt2.

(b) Again the translation is employed and the proof is similar to (a).

(c-d) Straightforward. I only need to show that  $K_i(T(A)) \Rightarrow K_i(A)$ . Assume  $K_i(T(A))$

then from KT1  $T(K_i(A))$  and by the T-T axiom for truth I obtain the desired  $K_i(A)$ . ■

**Proof. Lemma 18.** Assume  $M, w \models Bel_i(A)$ . Then for all  $w'$  such that  $\mathcal{B}_i(w, w')$  I obtain  $M, w' \models A$  and  $M, w' \not\models \neg A$ . From the semantic condition WSC, there is at least one world  $w'$  such that  $\mathcal{B}_i(w, w') \wedge \mathcal{K}_i(w, w')$ , therefore  $M, w \not\models Know_i(\neg A)$ , and hence  $M, w \models \neg Know_i(\neg A)$  as required. ■

**Proof. Lemma 19.** Assume  $M, w \models Bel_i(A)$ , then for all  $w'$  such that  $\mathcal{B}_i(w, w')$ ,  $M, w' \models A$  (i). Suppose  $\mathcal{K}_i(w, w')$  and  $\mathcal{B}_i(w', w'')$ , then I obtain  $M, w' \models Bel_i(A)$  and  $M, w \models Know_i(Bel_i(A))$ . From the semantic condition C1 I also get  $\mathcal{B}_i(w, w'')$  and thus  $M, w \models Bel_i(A)$  and  $M, w'' \models A$  as required by (i). ■

**Proof. Lemma 20.** Assume  $M, w \models Bel_i(A)$ , then for all  $w'$  such that  $\mathcal{B}_i(w, w')$ ,  $M, w \models A$  (i). Suppose  $\mathcal{K}_i(w', w'')$ , it now follows that  $M, w'' \models A$  and  $M, w' \models Know_i(A)$  and also  $M, w \models Bel_i(Know_i(A))$ . From the semantic condition C2 I obtain  $\mathcal{B}_i(w, w'')$  and thus  $M, w \models Bel_i(A)$  and  $M, w'' \models A$  as required by (i). ■

**Proof. Lemma 21.** Assume  $M, w \models \neg Bel_i(A)$ , then for some  $w'$  such that  $\mathcal{B}_i(w, w')$ ,  $M, w' \models \neg A$ . Assume that there is a  $w''$  such that  $\mathcal{B}_i(w, w'')$  then from the semantic condition C3 I obtain  $\mathcal{K}_i(w, w'')$  and thus  $M, w' \models \neg Know_i(A)$ . It now follows that  $M, w \models Bel_i(\neg Know_i(A))$  ■

**Proof. Proposition 22.1.** The  $\Rightarrow$  direction is already known. For the other direction  $\Leftarrow$ , assume  $Know_i(Bel_i(A))$ , then from the T-axiom for knowledge  $Know_i(Bel_i(A)) \Rightarrow Bel_i(A)$  and by an application of modus ponens  $Bel_i(\phi)$ .

2. The  $\Rightarrow$  is obvious, for  $\Leftarrow$  assume  $Know_i(A)$ . Then from the T-axiom for knowledge  $Know_i(A) \Rightarrow A$ . From the necessitation rule for belief I obtain  $Bel_i(Know_i(A) \Rightarrow A)$ . From the K-axiom for belief (distribution of belief over implication) it follows that  $Bel_i(Know_i(A)) \Rightarrow Bel_i(A)$ .

3. For  $\Rightarrow$ , assume  $\neg Bel_i(A)$ . From the S5 property of knowledge  $\psi \Rightarrow Know_i \neg Know_i \neg \psi$  and for  $\psi = \neg Bel_i(\phi)$  I obtain  $\neg Bel_i(\phi) \Rightarrow Know_i \neg Know_i \neg (\neg Bel_i(\phi))$ . As a result  $\neg Bel_i(\phi) \Rightarrow Know_i \neg Know_i (Bel_i(\phi))$  and by modus ponens  $Know_i \neg Know_i (Bel_i(\phi))$  (\*). From (1) of this proof I have that  $Bel_i(A) \Leftrightarrow Know_i(Bel_i(A))$  is a theorem of the logic and thus by negating both parts  $\neg Bel_i(A) \Leftrightarrow \neg Know_i(Bel_i(A))$ . By necessitation for knowledge  $Know_i(\neg Bel_i(A) \Leftrightarrow \neg Know_i(Bel_i(A)))$  and by k-K axiom I obtain  $Know_i(\neg Bel_i(A)) \Leftrightarrow Know_i(\neg Know_i(Bel_i(A)))$ . From this and (\*),  $Know_i(\neg Know_i(Bel_i(\phi))) \Rightarrow Know_i(\neg Bel_i(\phi))$  follows and by modus ponens  $Know_i(\neg Bel_i(\phi))$ . For the other direction  $\Leftarrow$ , assume that  $Know_i(\neg Bel_i(A))$ , then from the T-axiom of knowledge  $Know_i(\neg Bel_i(A)) \Rightarrow \neg Bel_i(A)$ , and by modus ponens  $\neg Bel_i(A)$ .

4. The  $\Rightarrow$  is obvious, for  $\Leftarrow$  assume  $Bel_i(\neg Know_i(A))$ . Then from the principle  $Bel_i(\neg \psi) \Rightarrow \neg Bel_i(\psi)$  and for  $\psi = Know_i(A)$  it follows that  $Bel_i(\neg Know_i(A)) \Rightarrow \neg Bel_i(Know_i(A))$ . Now by modus ponens  $\neg Bel_i(Know_i(A))$  (\*). I already know that  $Bel_i(Know_i(A)) \Leftrightarrow Bel_i(A)$  from (2) of this proposition. By negating both parts  $\neg Bel_i(Know_i(A)) \Leftrightarrow \neg Bel_i(A)$ . From this and (\*) it follows that  $Bel_i(\neg Know_i(A)) \Rightarrow \neg Bel_i(A)$  and by an application of modus ponens  $\neg Bel_i(A)$  ■

**Proof. Lemma 23.** Assume  $M, w \models Bel_i(A)$ . Then for all  $w'$  such that  $\mathcal{B}_i(w, w')$  I obtain  $M, w' \models A$ . From the semantic condition SC it follows that  $\mathcal{B}_i(w, w') \Rightarrow \mathcal{K}_i(w, w')$  and therefore  $M, w \models Know_i(A)$ . ■

**Proof. Theorem 24.** 1. Assume  $Know_i(A)$ . From the Skb axiom  $Know_i(A) \Rightarrow Bel_i(A)$  and by modus ponens  $Bel_i(A)$ . From the bt2 axiom  $Bel_i(A) \Rightarrow Bel_i(T(A))$  and from modus ponens I obtain the required  $Bel_i(T(A))$ .

2. Assume  $Know_i(A)$ . From the Skb axiom  $Know_i(A) \Rightarrow Bel_i(A)$  and by modus ponens  $Bel_i(A)$ . From the bt2 axiom  $Bel_i(A) \Rightarrow Bel_i(T(A))$  and from modus ponens  $Bel_i(T(A))$  is obtained. From the axiom bt1  $Bel_i(T(A)) \Rightarrow T(Bel_i(A))$  and by modus ponens again the required  $T(Bel_i(A))$  is obtained. ■

**Proof. Proposition 26.** 1. Similar to the proof of (1) of Proposition 22.

2. For the  $\Rightarrow$  direction, assume  $Know_i(A)$ . Then, from the positive introspection axiom for knowledge  $Know_i(A) \Leftrightarrow Know_i(Know_i(A))$  and by modus ponens  $Know_i(Know_i(A))$ . From the connection axiom of knowledge and belief I now obtain  $Know_i(Know_i(A)) \Leftrightarrow Bel_i(Know_i(A))$  and by another application of modus ponens it follows that  $Bel_i(Know_i(A))$ . For the  $\Leftarrow$  direction, assume  $\neg Know_i(A)$  (\*). By the negative introspection axiom for knowledge it is the case that  $\neg Know_i(A) \Rightarrow Know_i(\neg Know_i(A))$ , and by modus ponens  $Know_i(\neg Know_i(A))$ . By the connection axiom of knowledge and belief  $Know_i(\phi) \Rightarrow Bel_i(\phi)$  and for  $\phi = \neg Know_i(A)$  I obtain  $Know_i(\neg Know_i(A)) \Rightarrow Bel_i(\neg Know_i(A))$  and by modus ponens  $Bel_i(\neg Know_i(A))$ . From the principle  $Bel_i(\neg\psi) \Rightarrow \neg Bel_i(\psi)$ ,  $Bel_i(\neg Know_i(A)) \Rightarrow \neg Bel_i(Know_i(A))$  is obtained and by modus ponens I obtain  $\neg Bel_i(Know_i(A))$ . Thus,  $\vdash \neg Know_i(A) \Rightarrow \neg Bel_i(Know_i(A))$  and by contraposition  $Bel_i(Know_i(A)) \Rightarrow Know_i(A)$ .

3. Similar to the proof of (3) of Proposition 22.

4. For the  $\Rightarrow$  direction, assume  $\neg Know_i(A)$ . Then, by the negative introspection axiom for knowledge  $\neg Know_i(A) \Rightarrow Know_i(\neg Know_i(A))$ , and by modus ponens  $Know_i(\neg Know_i(A))$ . From the Skb axiom  $Know_i(\phi) \Rightarrow Bel_i(\phi)$  for  $\phi = \neg Know_i(A)$  I obtain  $Know_i(\neg Know_i(A)) \Rightarrow Bel_i(\neg Know_i(A))$  and thus by modus ponens  $Bel_i(\neg Know_i(A))$  and it follows that  $\neg Know_i(A) \Rightarrow Bel_i(\neg Know_i(A))$  as required. For the  $\Leftarrow$  direction, assume  $Know_i(A)$ . By the K-5 axiom  $Know_i(A) \Rightarrow Know_i(Know_i(A))$  and by modus ponens  $Know_i(Know_i(A))$ . From the theorem  $Know_i(\phi) \Rightarrow \neg Bel_i(\neg\phi)$  and  $\phi = Know_i(A)$  I obtain  $Know_i(Know_i(A)) \Rightarrow \neg Bel_i(\neg Know_i(A))$  and now by modus ponens  $\neg Bel_i(\neg Know_i(A))$ . It now follows that  $Know_i(A) \Rightarrow \neg Bel_i(\neg Know_i(A))$  and by contraposition  $Bel_i(\neg Know_i(A)) \Rightarrow \neg Know_i(A)$  is obtained.

5. Suppose  $Know_i(A)$ . Then, from the connection axiom  $Know_i(A) \Rightarrow Bel_i(A)$  and by modus ponens  $Bel_i(A)$ . From (1) of this proposition it is known that  $Bel_i(A) \Rightarrow Know_i(Bel_i(A))$  and by modus ponens  $Know_i(Bel_i(A))$ .

6. Suppose  $Know_i(A)$ . Then, from the connection axiom I obtain  $Know_i(A) \Rightarrow Bel_i(A)$  and by modus ponens  $Bel_i(A)$ . From the positive introspection axiom for belief  $Bel_i(A) \Rightarrow Bel_i(Bel_i(A))$  and by modus ponens  $Bel_i(Bel_i(A))$ . ■

**Proof. Lemma 27.** a) Assume  $M, w \models \neg Know_i(A)$ , then for some  $w'$  such that  $\mathcal{K}_i(w, w')$ ,  $M, w' \models \neg A$ . Assume that there is a  $w''$  such that  $\mathcal{B}_i(w, w'')$ , then from the semantic condition C4  $\mathcal{K}_i(w, w' w'')$  and thus  $M, w' \models \neg Know_i(A)$ . It now follows that  $M, w \models Bel_i(\neg Know_i(A))$ .

b) Assume  $M, w \models \neg Bel_i(A)$ , then for some  $w'$  such that  $\mathcal{B}_i(w, w')$ ,  $M, w' \models \neg A$ . Assume that there is a  $w''$  such that  $\mathcal{K}_i(w, w'')$ , then from the semantic condition C5  $\mathcal{B}_i(w, w' w'')$  and thus  $M, w' \models \neg Bel_i(A)$  and it now follows that  $M, w \models Know_i(\neg Bel_i(A))$ . ■

**Proof. Proposition 28.** The proofs of (1), (3-6) are similar to (1), (3-6) of Proposition 26, while the proof for (2) is similar to (2) of Proposition 22. ■

**Proof. Proposition 29.** 1. Assume  $B_i(A)$  then by employing the translation  $Bel_i(T(A))$ . From the modal axiom connecting knowledge and belief wkb  $Bel_i(T(A)) \Rightarrow \neg Know_i(\neg T(A))$  and by modus ponens  $\neg Know_i(T(\neg A))$ . Through the translation  $\neg K_i(\neg A)$  is obtained.

2. Assume  $B_i(A)$  then by the translation  $Bel_i(T(A))$ . From the modal axiom con-

necting knowledge and belief kb1  $Bel_i(T(A)) \Rightarrow Know_i(Bel_i(T(A)))$  and by modus ponens  $Know_i(Bel_i(T(A)))$ . From the kt2 axiom  $Know_i(Bel_i(T(A))) \Rightarrow Know_i(T(Bel_i(T(A))))$  is obtained and by modus ponens  $Know_i(T(Bel_i(T(A))))$ . By using the translation mechanism  $K_i(B_i(A))$ .

3. Assume  $B_i(A)$ . Via the translation  $Bel_i(T(A))$ . By the modal axiom kb2  $Bel_i(T(A)) \Rightarrow Bel_i(Know_i(T(A)))$  and by modus ponens  $Bel_i(Know_i(T(A)))$ . By bt2  $Bel_i(Know_i(T(A))) \Rightarrow Bel_i(T(Know_i(T(A))))$  and by modus ponens  $Bel_i(T(Know_i(T(A))))$ . By using the translation  $B_i(K_i(A))$ .

4. Assume  $\neg B_i(A)$  then by employing the translation  $\neg Bel_i(T(A))$ . From the modal axiom connecting knowledge and belief kb3  $\neg Bel_i(T(A)) \Rightarrow Bel_i(\neg Know_i(T(A)))$  is obtained and by modus ponens  $Bel_i(\neg Know_i(T(A)))$ . From the bt2 axiom it follows that  $Bel_i(\neg Know_i(T(A))) \Rightarrow Bel_i(T(\neg Know_i(T(A))))$  and by an application of modus ponens  $Bel_i(T(\neg Know_i(T(A))))$ . By using the translation the desired  $B_i(\neg K_i(A))$  is obtained. ■

**Proof. Proposition 30.** 1.  $\Rightarrow$  obvious from KB1 axiom. For the other direction  $\Leftarrow$  assume  $K_i(B_i(A))$ , then by the K-T axiom  $K_i(B_i(A)) \Rightarrow B_i(A)$  and by an application of modus ponens  $B_i(A)$ .

2. Assume  $K_i(A)$ , then by the K-4 axiom  $K_i(A) \Rightarrow K_i(K_i(A))$  and by modus ponens  $K_i(K_i(A))$ . From the theorem  $K_i(A) \Rightarrow \neg B_i(\neg A)$  it now follows that  $K_i(K_i(A)) \Rightarrow \neg B_i(\neg K_i(A))$  and by modus ponens  $\neg B_i(\neg K_i(A))$ . Thus,  $\vdash K_i(A) \Rightarrow \neg B_i(\neg K_i(A))$  and by contraposition  $\vdash B_i(\neg K_i(A)) \Rightarrow \neg K_i(A)$ . ■

**Proof. Proposition 31.** 1. Assume  $K_i(A)$ , then according to the translation  $Know_i(T(A))$ . From this and the strong connection axiom Skb  $Bel_i(T(A))$ , and by employing the translation mechanism  $B_i(A)$  as required.

2. Assume  $B_i(A)$ , then through the translation  $Bel_i(T(A))$ . By the modal axiom kb1  $Bel_i(T(A)) \Rightarrow Know_i(Bel_i(T(A)))$  and by modus ponens  $Know_i(Bel_i(T(A)))$ . By kt2  $Know_i(Bel_i(T(A))) \Rightarrow Know_i(T(Bel_i(T(A))))$  and by modus ponens  $Know_i(T(Bel_i(T(A))))$ . By using the translation  $K_i(B_i(A))$ . ■

**Proof. Proposition 32.** 1. Similar to (1) of Proposition 30.

2. Assume  $K_i(A)$ , then by the K-4 axiom  $K_i(A) \Rightarrow K_i(K_i(A))$  and by modus ponens  $K_i(K_i(A))$ . From the connection axiom SKB I now have  $K_i(K_i(A)) \Rightarrow B_i(K_i(A))$  and by modus ponens again  $B_i(K_i(A))$ .

3. Suppose  $K_i(A)$ , then from the connection axiom SKB  $K_i(A) \Rightarrow B_i(A)$  and by modus ponens  $B_i(A)$ . Now from (4) of this proposition it is also known that  $B_i(A) \Rightarrow K_i(B_i(A))$  and thus by modus ponens  $K_i(B_i(A))$ .

4. Suppose  $K_i(A)$ , then from the connection axiom SKB  $K_i(A) \Rightarrow B_i(A)$  and by modus ponens  $B_i(A)$ . From the B-4 axiom  $B_i(A) \Rightarrow B_i(B_i(A))$  and thus by modus ponens  $B_i(B_i(A))$ . ■

**Proof. Proposition 33.** 1. From the SKB axiom  $K_i(A) \Rightarrow B_i(A)$  and by modus ponens  $B_i(A)$ . From axiom BT2  $B_i(A) \Rightarrow B_i(T(A))$  and by modus ponens the required  $B_i(T(A))$  is obtained.

2. From the SKB axiom  $K_i(A) \Rightarrow B_i(A)$  and by modus ponens  $B_i(A)$ . From axiom BT2 I obtain  $B_i(A) \Rightarrow B_i(T(A))$  and by modus ponens  $B_i(T(A))$ . From axiom BT1  $B_i(T(A)) \Rightarrow T(B_i(A))$  and by modus ponens again the required  $T(B_i(A))$  is obtained. ■

**Proof. Proposition 34.** For (1-2) the proof is similar to (1-2) of Proposition 31.

3. Assume  $B_i(A)$ , then by the translation  $Bel_i(T(A))$ . From the modal axiom kb2  $Bel_i(T(A)) \Rightarrow Bel_i(Know_i(T(A)))$  and from modus ponens  $Bel_i(Know_i(T(A)))$ . By bt2

Code	Wff	Conditions	D	D	D	T	T	T	S4	S4	S4	S5	S5	S5
			D	S4	S5	D	S4	S5	D	S4	S5	D	S4	S5
w1	$Bel(A) \Rightarrow \neg Know(\neg A)$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w2	$Bel(A) \Rightarrow Know(Bel(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w3	$Know(Bel(A)) \Rightarrow Bel(A)$	k-T				✓	✓	✓	✓	✓	✓	✓	✓	✓
w4	$\neg Bel(A) \Rightarrow Know(\neg Bel(A))$	k-5 & $\Leftrightarrow$ of w2										✓	✓	✓
w5	$Know(\neg Bel(A)) \Rightarrow \neg Bel(A)$	k-T				✓	✓	✓	✓	✓	✓	✓	✓	✓
w6	$Bel(A) \Rightarrow Bel(Know(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w7	$Bel(Know(A)) \Rightarrow Bel(A)$	k-T				✓	✓	✓	✓	✓	✓	✓	✓	✓
w8	$\neg Bel(A) \Rightarrow Bel(\neg Know(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w9	$Bel(\neg Know(A)) \Rightarrow \neg Bel(A)$	$\Leftrightarrow$ of w6				✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 1: Modal systems of knowledge and belief and Weak connection axiomatisation.

$Bel_i(Know_i(T(A))) \Rightarrow Bel_i(T(Know_i(T(A))))$  and by modus ponens  $Bel_i(T(Know_i(T(A))))$ . By using the translation mechanism again  $B_i(K_i(A))$ .

4. Assume  $\neg K(A)$ , then by the translation I obtain  $\neg Know(T(A))$ . Since  $\neg Know_i(T(A)) \Rightarrow Bel_i(\neg Know_i(T(A)))$  is a theorem by modus ponens  $Bel_i(\neg Know_i(T(A)))$ . From the bt2 axiom  $Bel_i(\neg Know_i(T(A))) \Rightarrow Bel_i(T(\neg Know_i(T(A))))$  and by an application of modus ponens  $Bel_i(T(\neg Know_i(T(A))))$  and by the translation  $B_i(\neg K_i(A))$ . ■

**Proof. Proposition 35.** For (1-4) the proof is similar to that of Proposition 32. For (5), the  $\Rightarrow$  direction is obvious. For the other direction  $\Leftarrow$ , the proof is similar to (2) of Proposition 30. ■

## A.2 Weaker systems

### A.2.1 Weaker Systems for modal knowledge and belief

Since space constraints do not allow the presentation of all the weaker systems for modal knowledge and belief, in this section a synopsis of the properties in weaker systems is provided. For each connection axiomatisation (weak, strong s1 and strong s2) a table is provided that contains a number of coded properties and their satisfaction in a combined modal system of knowledge and belief. The first letter indicates the strongest axiom for knowledge while the second the strongest axiom for belief. Thus T, S5 indicates that the combined system is that of T for knowledge and KD45 for belief. These tables will provide a guide to the interested reader, who may want weaker axiomatisations of knowledge and belief. Accordingly, there are three families of combined modal logics.

i) w-TKB<sup>m</sup>[X,Y,Z]. These are the logics that are connected via the weak type of connection. TKB<sup>m</sup> is a logic of truth and modal knowledge and belief as it was defined in Section 6.2.

ii) s1-TKB<sup>m</sup>[X,Y,Z]. These are the logics that are connected via the strong s1 type of connection.

iii) s2-TKB<sup>m</sup>[X,Y,Z]. These are the logics that are connected via the strong s2 type of connection. Although TKB<sup>m</sup> is as above, Y does not include the k-5 axiom (negative introspection for knowledge).

Code	Wff	Conditions	D	D	D	T	T	T	S4	S4	S4	S5	S5	S5
			D	S4	S5	D	S4	S5	D	S4	S5	D	S4	S5
s1.1	$Know(A) \Rightarrow Bel(A)$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
s1.2	$Know(A) \Rightarrow \neg Bel(\neg A)$	s1.1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
s1.3	$Bel(A) \Rightarrow Know(Bel(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
s1.4	$Know(Bel(A)) \Rightarrow Bel(A)$	k-T				✓	✓	✓	✓	✓	✓	✓	✓	✓
s1.5	$\neg Bel(A) \Rightarrow Know(\neg Bel(A))$	k-5 & $\Leftrightarrow$ of s1.3										✓	✓	✓
s1.6	$Know(\neg Bel(A)) \Rightarrow \neg Bel(A)$	k-T				✓	✓	✓	✓	✓	✓	✓	✓	✓
s1.7	$Know(A) \Rightarrow Bel(Know(A))$	k-4							✓	✓	✓	✓	✓	✓
s1.8	$Bel(Know(A)) \Rightarrow Know(A)$	k-5										✓	✓	✓
s1.9	$\neg Know(A) \Rightarrow Bel(\neg Know(A))$	k-5										✓	✓	✓
s1.10	$Bel(\neg Know(A)) \Rightarrow \neg Know(A)$	k-4							✓	✓	✓	✓	✓	✓
s1.11	$Know(A) \Rightarrow Know(Bel(A))$	s1.3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
s1.12	$Know(A) \Rightarrow Bel(Bel(A))$	b-4		✓	✓		✓	✓		✓	✓			✓

Table 2: Modal systems of knowledge and belief and Strong s1 connection axiomatisation.

Code	Wff	Conditions	D	D	D	T	T	T	S4	S4	S4
			D	S4	S5	D	S4	S5	D	S4	S5
s2.1	$Know(A) \Rightarrow Bel(A)$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓
s2.2	$Know(A) \Rightarrow \neg Bel(\neg A)$	s2.1	✓	✓	✓	✓	✓	✓	✓	✓	✓
s2.3	$Bel(A) \Rightarrow Know(Bel(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓
s2.4	$Know(Bel(A)) \Rightarrow Bel(A)$	k-T				✓	✓	✓	✓	✓	✓
s2.5	$Bel(A) \Rightarrow Bel(Know(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓
s2.6	$Bel(Know(A)) \Rightarrow Bel(A)$	k-T				✓	✓	✓	✓	✓	✓
s2.7	$Know(A) \Rightarrow Know(Bel(A))$	s2.3	✓	✓	✓	✓	✓	✓	✓	✓	✓
s2.8	$Know(A) \Rightarrow Bel(Bel(A))$	b-4		✓	✓		✓	✓		✓	✓
s2.9	$\neg Know(A) \Rightarrow Bel(\neg Know(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓
s2.10	$Bel(\neg Know(A)) \Rightarrow \neg Know(A)$	k-4							✓	✓	✓
s2.11	$\neg Bel(A) \Rightarrow Know(\neg Bel(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓
s2.12	$Know(\neg Bel(A)) \Rightarrow \neg Bel(A)$	k-T				✓	✓	✓	✓	✓	✓

Table 3: Modal systems of knowledge and belief and Strong s2 connection axiomatisation.



Code	Wff	Conditions	D	D	T	T	S4	S4
			D	S4	D	S4	D	S4
W1	$B(A) \Rightarrow \neg K(\neg A)$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
W2	$B(A) \Rightarrow K(B(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
W3	$B(A) \Rightarrow B(K(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
W4	$\neg B(A) \Rightarrow B(\neg K(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
W5	$K(A) \Rightarrow \neg B(\neg A)$	W1	✓	✓	✓	✓	✓	✓
W6	$B(A) \Leftrightarrow K(B(A))$	W2 and K-T			✓	✓	✓	✓
W7	$B(\neg K(A)) \Rightarrow \neg K(A)$	K-4 and W1					✓	✓

Table 4: Syntactical systems of knowledge and belief and Weak connection axiomatisation.

Code	Wff	Conditions	D	D	T	T	S4	S4
			D	S4	D	S4	D	S4
S1.1	$K(A) \Rightarrow B(A)$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
S1.2	$B(A) \Rightarrow K(B(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
S1.3	$B(A) \Rightarrow \neg K(\neg A)$	S1.1	✓	✓	✓	✓	✓	✓
S1.4	$B(A) \Leftrightarrow K(B(A))$	S1.2 and K-T			✓	✓	✓	✓
S1.5	$K(A) \Rightarrow B(K(A))$	K-4 and S1.1					✓	✓
S1.6	$K(A) \Rightarrow K(B(A))$	S1.1 and S1.2					✓	✓
S1.7	$K(A) \Rightarrow B(B(A))$	S1.1 and B-4		✓		✓		✓

Table 5: Syntactical systems of knowledge and belief and Strong S1 connection axiomatisation.

### A.2.2 Weaker systems for syntactic knowledge and belief

The connection axiomatisations for the logic of truth and syntactic modalities were again investigated in the context of the strongest system that does not allow knowledge and belief to collapse. Intermediate results can be obtained for weaker systems of knowledge and belief, as in the modal case. In this section a synopsis of the properties in the weaker systems is provided. For each connection axiomatisation (weak, strong S1 and strong S2) a table is provided that contains a number of coded properties and their satisfaction in a combined syntactical system of knowledge and belief. The first letter indicates the strongest axiom for knowledge while the second the strongest axiom for belief. Thus T, S4 indicates that the combined system is that of T for knowledge and KD4 for belief. There are three families of combined syntactical logics:

- i) W-STKB[X,Y,Z]: These are the logics that are connected via the weak connection axiomatisation. STKB is a logic of truth and syntactic knowledge and belief as it was defined in Section 8.
- ii) S1-STKB[X,Y,Z]: These are the logics that are connected via the strong S1 connection axiomatisation.
- iii) S2-STKB[X,Y,Z]: These are the logics that are connected via the strong S2 connection axiomatisation.

**Proposition 36** *a) All theorems of a logic W-STKB[X, Y, Z] are theorems of a corresponding system w-TKB<sup>m</sup>[X, Y, Z].*

Code	Wff	Conditions	D	D	T	T	S4	S4
			D	S4	D	S4	D	S4
S2.1	$K(A) \Rightarrow B(A)$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
S2.2	$B(A) \Rightarrow K(B(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
S2.3	$B(A) \Rightarrow B(K(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
S2.4	$\neg K(A) \Rightarrow B(\neg K(A))$	<i>Axiom</i>	✓	✓	✓	✓	✓	✓
S2.5	$B(A) \Rightarrow \neg K(\neg A)$	S2.1	✓	✓	✓	✓	✓	✓
S2.6	$B(A) \Leftrightarrow K(B(A))$	S2.1 and K-T			✓	✓	✓	✓
S2.7	$K(A) \Rightarrow B(K(A))$	S2.1 and K-4					✓	✓
S2.8	$K(A) \Rightarrow K(B(A))$	S2.1 and K-4					✓	✓
S2.9	$K(A) \Rightarrow B(B(A))$	S2.1 and B-4		✓		✓		✓
S2.10	$\neg K(A) \Leftrightarrow B(\neg K(A))$	S2.4 and K-4					✓	✓

Table 6: Syntactical systems of knowledge and belief and Strong S2 connection axiomatisation.

b) All theorems of a logic  $S1\text{-STKB}[X, Y, Z]$  are theorems of a corresponding system  $s1\text{-TKB}^m[X, Y, Z]$ .

c) All theorems of a logic  $S2\text{-STKB}[X, Y, Z]$  are theorems of a corresponding system  $s2\text{-TKB}^m[X, Y, Z]$ .

**Proof.** The proof is established by using the translation from the syntactical system to the respective modal system. ■