A Computationally Grounded Logic of Knowledge, Belief and Certainty

Kaile Su, Abdul Sattar
Institute for Integrated and Intelligent Systems
Griffith University
Brisbane, Australia
{k.su, a.sattar}@grif th.edu.au

Guido Governatori
School of ITEE
The University of Queensland
Brisbane, Australia
guido@itee.uq.edu.au

Qingliang Chen
Dept. of Computer Science
Sun Yat-sen University
Guangzhou, P.R. China
tsingliangchen@yahoo.com

ABSTRACT

This paper presents a logic of knowledge, belief and certainty, which allows us to explicitly express the knowledge, belief and certainty of an agent. A computationally grounded model, called interpreted KBC systems, is given for interpreting this logic. The relationships between knowledge, belief and certainty are explored. In particular, certainty entails belief; and to the agent what it is certain of appears to be the knowledge. To formalize those agents that are able to introspect their own belief and certainty, we identify a subclass of interpreted KBC systems, called *introspective KBC systems*. We provide sound and complete axiomatizations for the logics. We show that the validity problem for the interpreted KBC systems is PSPACE-complete, and the same problem for introspective KBC systems is co-NP complete, thus no harder than that of the propositional logic.

Categories and Subject Descriptors

F.3.1 [**Logics and Meanings of Programs**]: Specifying and Verifying and Reasoning about Programs—*Mechanical verification, Specification techniques*; I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Intelligent agents, Multiagent systems*

General Terms

Theory

Keywords

modal logic, interpreted system, computationally grounded model

1. INTRODUCTION

Philosophers have been interested in the notions of knowledge and belief for a long time, and with the advent of agent oriented computing these notions are of paramount importance for the formalization of autonomous agents. Modal logic has proved to be a suitable formalism for the understanding of the formal properties of knowledge, belief and many other mental attitudes. Furthermore, the formal study of these notions should involve the development of semantics to give a proper and unambiguous meaning to the notions involved, to compare different systems, and to evaluate different intuitions about their interpretations.

There are two main semantic approaches to formalizing agent systems via modal logics, the *possible worlds semantics* [10, 12] and the *interpreted system* model [6, 9, 8]. The first approach is very fruitful, including the well-known theory of intention [3] and the formalism of the belief-desire-intension paradigm [18]. The second, mainly due to Halpern and his colleagues [6, 9, 8], offers a very natural interpretation, in terms of the states of computer processes, to S5 epistemic logic. The advantage of the first approach is that internal mental attitudes of an agent, such as beliefs and goals, can be characterized conveniently with a model theoretic feature in terms of the belief and goal accessibility relations. On the contrast, the salient point of second approach is that we are able to associate the system with a computer program and formulas can be understood as properties of program computations. In this sense, the interpreted system model is *computationally grounded* [22].

An extension to the interpreted system model has been given by the computationally grounded logic $\mathcal{VSK}$ [23], which enable us to represent what is *visible* in the environment to individual agents, what these agents actually *perceive*, and what the agents actually *know* about the environment. The semantics of the logic is given in terms of a general model of multi-agent systems called the *agent-environment system*. Notice that all three modalities $\mathcal{V}$, $\mathcal{S}$ and $\mathcal{K}$ in the logic as well as knowledge operator in Halpern's epistemic logic can be regarded as external notions to the agent. Consider, for example, formula $\mathcal{S}\varphi$ in $\mathcal{VSK}$ logic, which means that *the percept received by the agent carries the information* φ; thus, an agent may have no way to know whether $\mathcal{S}\varphi$ holds.

The aim of this paper is to develop a general formalism to represent and reason about the knowledge, belief and certainty of an agent. The relationships between knowledge, belief and certainty are explored. In particular, 'certainty' is a kind of belief and it might be false, but the agent or human who feels certain of a proposition is not aware that the proposition might be false. Also, 'certainty' is closely related to 'knowledge', and to the agent itself what it is certain of appears to be knowledge.

Syntactically, the formalism we develop here is a propositional multimodal logic, containing unary modal operators K, B, and C. A formula $K\varphi$ means that the agent knows the information φ; $B\varphi$

This work was supported by the Australian Research Council grant DP0452628, and partially by the NSFC grants 60496327, 10410638 and 60473004.

AAMAS'05, July 25›29, 2005, Utrecht, Netherlands.

means that the agent believes the information φ; and $C\varphi$ means that the agent is certain of the information φ. The semantics of the logic is given with respect to a simple, natural, and computationally grounded model of agents and their environments, which is called *interpreted KBC system* model.

A key feature of the interpreted KBC system model is that it extends the interpreted system model, and in particular the agent-environment system of $\mathcal{VSK}$ [23], in the following two aspects:

We assume that the sensors of an agent may become inaccurate and the agent may not be aware of this. Thus, the visible part of the environment may differ from the percept received by the agent.

We suppose that an agent can access some external information sources and get its belief about those parts of environment that are invisible to the agent.

We believe that the interpreted KBC system model is not only appropriate for our logic, but also useful in the design and analysis of agent systems because of the modularity of both external and internal parts of an interpreted KBC system. In particular it is possible to identify a subclass of interpreted KBC systems, called *introspective KBC systems*, that represents agents' ability to introspect their own knowledge, belief, and certainty.

The significance of our logic is as follows. First, it extends epistemic logic with the belief modality B and the certainty modality C and maintains the computational grounding of the original formalism. Secondly, the modal operators B and C in our logic need not satisfy the S5 axioms as those in $\mathcal{VSK}$ logic do. Finally, we shed light on the relationships between knowledge, belief and certainty by using the concrete model of extended interpreted systems presented in this paper.

The paper is organized as follows. In the next section, we introduce the computational model upon which our work is based. Then, in Section 3 we present our computationally grounded logics KBC and introspective KBC. In Section 4.1 we address the issue of completeness of the two logics and we show the equivalence between appropriate Kripke semantics and interpreted (introspective) KBC systems. Then in Section 4.2 we investigate the computational complexity of the logics and we prove by using the well-known tableau method [13, 1], the validity problem in KBC is in PSPACE, and the same problem for introspective KBC systems is co-NP complete, which is no harder than that of the propositional logic. Finally, in Section 5, we discuss some related work and identify future extensions.

2. A COMPUTATIONAL MODEL

The systems we are modelling, called *KBC systems*, consist of an environment and an agent who has knowledge, belief and certainty about the environment. Both the environment and the agent are in some states at any point of time, which are referred to as the environment state and the agent's internal state respectively, in order to distinguish them from the state of the systems, i.e., the *global* state. We divide the environment state into two parts: the visible and invisible parts. As for the agent's internal state, we distinguish what the agent sees or perceives about the visible part of the environment, from what it conjectures or believes the invisible part of the environment to be. We give a visual representation of this framework in Figure 1.

Formally, an *environment state* Es is a pair (s_{vis}, s_{inv}) of a *visible part* s_{vis} and an *invisible part* s_{inv}. An *internal state* Is is a pair (s_{per}, S_{pls}), where s_{per} is the agent's perception of the visible part of the environment state, and S_{pls} is a set of plausible invisible parts of the environment state that the agent thinks possible. A *global state* is a pair of an environment state and an internal state; in other words, it is a 4-tuple $(s_{vis}, s_{inv}, s_{per}, S_{pls})$ of the visible part of environment, the invisible part of environment, the agent's perception, and the plausible set.

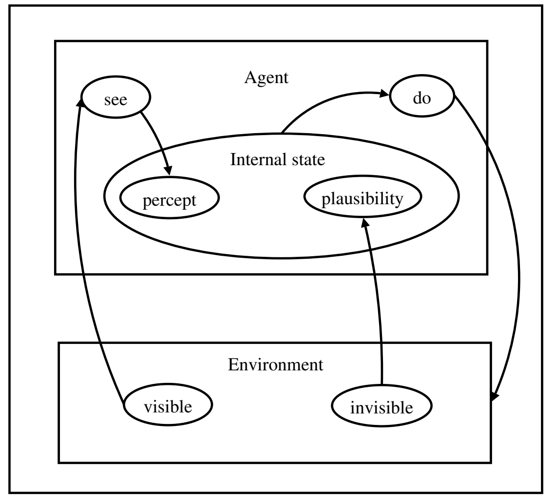

Figure 1: A Computational Model

Given a global state $s = (s_{vis}, s_{inv}, s_{per}, S_{pls})$, we denote $s_{vis}, s_{inv}, s_{per}$ and S_{pls} by $Vis(s)$, $Inv(s)$, $Per(s)$ and $Pls(s)$, respectively.

Let V denote the set of visible parts of environment states, and I denote the set of invisible parts of environment states. We have that any environment state is in $V \quad I$, and any internal state is in $V \quad 2^I$. Therefore, any global state is in $V \quad I \quad V \quad 2^I$. We take $G \quad V \quad I \quad V \quad 2^I$ to be the set of *reachable global* states of the system. A *run* over G is a function from the time domain–the natural numbers in our case–to G. Thus, a run over G can be identified with a sequence of global states in G. We refer to a pair (r, m) consisting of a run r and time m as a *point*.

The idea of our computational model is that a run represents one possible computation of a KBC system and a KBC system may have a number of possible runs, so we say a KBC system is a set of runs.

We may assume that changes in the global state are caused by actions performed by the agent. Thus we can define the notion of a KBC system in terms of *actions*, where an action is a function from the set of global states into the set of global states, representing changes of the environment and changes in the belief of an agent about the invisible part of the environment. Given an environment state and internal state of the agent, there may be more than one possible actions that can be executed by the agent; thus, we may have multiple runs in a KBC system.

(0,1), clean	(1,1), dirty
(0,0), clean	(1,0), clean

Figure 2: Four Environment States with Both Visible and Invisible Parts

EXAMPLE 1. *Let us consider the scenario depicted in Figure 2: There are four squares in a 2x2 square, one of them, say, the top right one is dirty. Suppose that a robot stands at one of the square, and believes that the right two squares are dirty. The robot can move horizontally or vertically and can move from one square to another neighboring square. However, the robot's moving-direction sensors is broken, and the robot may get uncorrected perception about its location. For instance, if the robot vertically move from square* $(0,0)$ *into square* $(0,1)$, *it may think it horizontally moved into square* $(1,0)$*; thus, the robot may confuse square* $(0,1)$ *with square* $(1,0)$. *In the same way, the robot can not correctly distinguish square* $(0,0)$ *from square* $(1,1)$. *Generally speaking, a global state can be represented as*

$$((x_e, y_e), z, (x_a, y_a), Z)$$

where

> x_e, y_e, z, x_a, y_a *are boolean value,* (x_e, y_e) *indicates which square the robot located at,* z *expresses whether the square is dirty,* (x_a, y_a) *is used for what the robot perceives about its location;*
>
> Z *is a subset of* $\{0,1\}$, *and we have that if* $Z = \{1\}$, *then the robot believes that the square is dirty, if* $Z = \{0\}$, *then the robot believes that the square is clean, if* $Z = \{0,1\}$, *then the robot has contradicting belief , and finally, if* $Z = \{\}$, *then the robot has no idea about whether the square is clean.*

According to our discussion above, we have several constraints on these variables:

> $x_e \wedge y_e \Leftrightarrow z$ *holds, which indicates that only square* $(1,1)$ *is dirty;*
>
> $(x_e \Leftrightarrow y_e) \Leftrightarrow (x_a \Leftrightarrow y_a)$ *holds which means that the robot is able to distinguish two neighboring squares;*
>
> *if* y_a *is* 1, *then* Z *is* $\{1\}$, *this say that the robot believes that the right two squares are dirty.*

Now we define a KBC *system as the set of those runs where, given a global state* $g = ((x_e, y_e), z, (x_a, y_a), Z)$, *the square* (x'_e, y'_e) *of next global state* $g' = ((x'_e, y'_e), z', (x'_a, y'_a), Z')$ *neighbours to square* (x_e, y_e) *(that is,* $x_e = x'_e$ *and* $y_e \neq y'_e$, *or* $x_e \neq x'_e$ *and* $y_e = y'_e$*) and the next global state still satisfies the constraints above.*

In our model, environment states consist of the visible and invisible parts; whatever is not invisible must be visible. Moreover, the visible and invisible parts may be dependent on each other. In the above example, for the environment state $((x_e, y_e), z)$, the relationship between the visible part (x_e, y_e) and the invisible part z is represented as $x_e \wedge y_e \Leftrightarrow z$.

Assume that we have a set of primitive propositions, which we can think of as describing basic facts about the system. An *interpreted* KBC *system* $\mathcal{I}$ consists of a pair $(\mathcal{R},\ )$, where $\mathcal{R}$ is a set of runs over a set of global states and is a valuation function, which assigns truth values to primitive propositions at each environment state Es in $\mathcal{R}$. Thus, for every $p \in$ and every point (r,m) in $\mathcal{R}$, we have $(Vis(r(m)), Inv(r(m)))(p) \in \{\mathbf{true}, \mathbf{false}\}$. For instance, for the KBC system in Example 1, we may take $\{left, top, dirty\}$ as the set of primitive propositions, representing the location of the square where the robot stand in, and whether the square is clean or dirty. Clearly, for the environment state $Es = ((0,1), 0)$, we may naturally define $(Es)(left) = \mathbf{true}$, $(Es)(top) = \mathbf{true}$ and $(Es)(dirty) = \mathbf{false}$.

To define knowledge in interpreted KBC systems, we introduce an equivalence relation $_{vis}$ over the set of points: $(r,u)\ _{vis}\ (r',v)$ iff $Vis(r(u)) = Vis(r'(v))$.

If $(r,u)\ _{vis}\ (r',v)$, then we say that (r,u) and (r',v) are indistinguishable to the agent.

Interpreted KBC systems play the same role in our logic as what interpreted systems play in epistemic logic. We also call interpreted KBC systems generalized interpreted systems.

Finally, we remark that the states are used to define runs over time, which form the basis of the semantics for the logic presented in the next section. However, from the logical point of view, it is enough to consider accessibility relations because there is no dynamic aspect in this logic. After all, a set of runs, can be mapped in an accessibility relation, and vice versa. The main reasons we do this is to demonstrate how this logic could model the agent performing actions in the environment. Moreover, it is also interesting to extend this logic by incorporating some modalities of temporal logic.

3. A COMPUTATIONALLY GROUNDED LOGIC

In this section, we introduce a multimodal logic of knowledge, belief and certainty, called KBC logic, which enables us to represent the knowledge, belief, and certainty of an agent about some environment. The semantics of KBC logic is given in terms of the computational model presented above, and thus the logic is computationally grounded in this sense.

3.1 Syntax

Given a set of propositional atoms, the language of KBC logic is defined by the following BNF grammar:

$$\begin{aligned} \langle wff \rangle ::= \ & \text{any element of } \ \mid \\ & \neg\langle wff \rangle \mid \langle wff \rangle \wedge \langle wff \rangle \mid \\ & K\langle wff \rangle \mid B\langle wff \rangle \mid C\langle wff \rangle \end{aligned}$$

The modality K allows us to represent what is observable about an environment state. Formula $K\phi$ means that ϕ is observable about the environment. $C\phi$ means that the agent feels certain for the property ϕ. Intuitive meaning behind 'certainty' is that, to the agent, the facts of which he is certain appear to be knowledge [14]. Finally $B\phi$ says that the agent believes the property ϕ.

3.2 Semantics

We now proceed to interpret the KBC logic formulas in terms of interpreted KBC systems. In the following, we define the satisfaction relation $\models$ between a formula and a pair of interpreted environment-agent system and a point. Given an interpreted system $\mathcal{I} = (\mathcal{R},\ )$ and a point (r,u) in $\mathcal{I}$, we have that:

> $(\mathcal{I}, r, u) \models p$ iff $(Vis(r(u)), Inv(r(u)))(p) = \mathbf{true}$, where p is a primitive proposition in ;
>
> $(\mathcal{I}, r, u) \models \neg\varphi$ iff $(\mathcal{I}, r, u) \not\models \varphi$;
>
> $(\mathcal{I}, r, u) \models \varphi \wedge$ iff $(\mathcal{I}, r, u) \models \varphi$ and $(\mathcal{I}, r, u) \models$;
>
> $(\mathcal{I}, r, u) \models K\varphi$ iff $(\mathcal{I}, r', v) \models \varphi$ for those (r', v) such that $(r,u)\ _{vis}\ (r',v)$;
>
> $(\mathcal{I}, r, u) \models B\varphi$ iff $(\mathcal{I}, r', v) \models \varphi$ for those (r', v) such that $Vis(r'(v)) = Per(r(u))$ and $Inv(r'(v)) \in Pls(r(u))$;
>
> $(\mathcal{I}, r, u) \models C\varphi$ iff $(\mathcal{I}, r', v) \models \varphi$ for those (r', v) such that $Vis(r'(v)) = Per(r(u))$.

According to the definitions above, the agent knows φ at point (r,u) if and only if φ holds at those points with the same visible parts as that point. The agent is certain of φ means that φ holds at those points with the agent's current perception as the visible parts, whereas the agent believes φ means that φ holds at those points with the visible part equaling to the agent's current perception and with the invisible part plausible from the agent's point of view.

We say that a formula φ is valid in an interpreted KBC system $\mathcal{I}$, denoted by $\mathcal{I} \models \varphi$, if $(\mathcal{I}, r, u) \models \varphi$ holds for every point (r,u) in $\mathcal{I}$. We use $\models \varphi$ to denote that φ is valid in every interpreted KBC systems.

We notice that knowledge is an *external* notion–an agent is said to know φ if an impartial, omniscient observer would say that the agent's state carries the information φ. $K\varphi$ means that not only φ is true of the environment, but also the agent would be able to perceive φ if its sensor apparatus was good enough. Our interpretation of $C\varphi$ captures the intuition behind 'certainty' that, to the agent, the fact of which it feels certain appears to be knowledge. Thus, 'John is certain that' is equivalent to 'John is certain that John knows'. Certainty entails belief, but the agent gets its belief not only via what it perceives about the visible part of the environment state but also via what, it conjectures, the invisible part of the environment to be like.

3.3 Valid formulas of KBC logic

Let us consider what formulas are *valid* for every interpreted KBC systems.

PROPOSITION 2. *The following formulas are valid:*

knowledge
$K(\varphi \Rightarrow \quad) \Rightarrow (K\varphi \Rightarrow K \quad)$
$K\varphi \Rightarrow \varphi$
$K\varphi \Rightarrow KK\varphi$
$\neg K\varphi \Rightarrow K\neg K\varphi$

belief
$B(\varphi \Rightarrow \quad) \Rightarrow (B\varphi \Rightarrow B \quad)$

certainty
$C(\varphi \Rightarrow \quad) \Rightarrow (C\varphi \Rightarrow C \quad)$

certainty and knowledge
$C\varphi \Rightarrow CK\varphi$
$\neg C\varphi \Rightarrow C\neg K\varphi$

certainty and belief
$C\varphi \Rightarrow B\varphi$

Proof: We prove that $\models C\varphi \Rightarrow CK\varphi$ and $\models \neg C\varphi \Rightarrow C\neg K\varphi$; other cases are similar or easier. Given an interpreted system $\mathcal{I} = (\mathcal{R}, \quad)$ and a point (r,u) in $\mathcal{I}$, suppose that $(\mathcal{I}, r, u) \models C\varphi$. Then, for every point (r', v) with $Vis(r'(v)) = Per(r(u))$, we have that $(\mathcal{I}, r', v) \models \varphi$. To show $(\mathcal{I}, r, u) \models CK\varphi$, we must prove that, for every point (r_1, u_1) with $Vis(r_1(u_1)) = Per(r(u))$, we have that $(\mathcal{I}, r_1, u_1) \models K\varphi$. It suffices to prove that, for every point (r_2, u_2) with $Vis(r_1(u_1)) = Vis(r_2(u_2))$, $(\mathcal{I}, r_2, u_2) \models \varphi$ holds. However, we get $Vis(r_2(v_2)) = Per(r(u))$ from both $Vis(r_1(u_1)) = Vis(r_2(u_2))$ and $Vis(r_1(u_1)) = Per(r(u))$. Thus, $(\mathcal{I}, r_2, u_2) \models \varphi$ holds by the condition $(\mathcal{I}, r, u) \models C\varphi$. This proves that $\models C\varphi \Rightarrow CK\varphi$.

In order to prove that $\models \neg C\varphi \Rightarrow C\neg K\varphi$, we suppose $(\mathcal{I}, r, u) \models \neg C\varphi$. Then, for some point (r', v) with $Vis(r'(v)) = Per(r(u))$, we have that $(\mathcal{I}, r', v) \not\models \varphi$. To show $(\mathcal{I}, r, u) \models C\neg K\varphi$, we must prove that, for every point (r_1, u_1) with $Vis(r_1(u_1)) = Per(r(u))$, we have that $(\mathcal{I}, r_1, u_1) \models \neg K\varphi$. It suffices to prove that, there exists some point (r_2, u_2) with $Vis(r_1(u_1)) = Vis(r_2(u_2))$ such that $(\mathcal{I}, r_2, u_2) \not\models \varphi$ holds. However, we have that $Vis(r_1(u_1)) = Vis(r_2(u_2)) = Per(r(u))$ and $(\mathcal{I}, r', v) \not\models \varphi$. Hence, $(\mathcal{I}, r, u) \models C\neg K\varphi$. This completes the proof. ∎

To formalize agents' ability to introspect their own belief or certainty, we identify some subclasses of interpreted KBC systems as follows:

> We say that $\mathcal{I}$ is a *KC-introspective KBC system* if for every two points (r,u) and (r', v) in $\mathcal{R}$, $Vis(r(u)) = Vis((r'(v))$ implies $Per(r(u)) = Per(r'(v))$.
>
> We say that $\mathcal{I}$ is a *C-introspective KBC system* if for every (r,u) in $\mathcal{R}$ and every (r', v) in $\mathcal{R}$ such that $Vis(r'(v)) = Per(r(u))$, we have that $Per(r(u)) = Per(r'(v))$.
>
> We say that $\mathcal{I}$ is a *CB-introspective KBC system* if for every two points (r,u) and (r', v) in $\mathcal{R}$ such that $Vis(r'(v)) = Per(r(u))$, we have that $Pls(r(u)) = Pls(r'(v))$.
>
> We say that $\mathcal{I}$ is a *B-introspective KBC system* if for every (r,u) in $\mathcal{R}$ and every (r', v) in $\mathcal{R}$ such that $Vis(r'(v)) = Per(r(u))$ and $Inv(r'(v)) \in Pls(r(u))$, we have that $Pls(r(u)) = Pls(r'(v))$.
>
> Finally, we say that $\mathcal{I}$ is an *introspective KBC system* if $\mathcal{I}$ is a KC-introspective, C-introspective, and CB-introspective KBC system.

We notice that for a KC-introspective KBC system, the perception of an agent is determined by the visible part of the environment. The intuition behind the notion of the C-introspective KBC system, on the other hand, is that the agent is confident about its perception apparatus, that is, the agent believes that it has perfect perception. For a CB-introspective KBC system, the agent thinks its belief is determined by the visible part of the environment, and the visible part of the environment in the agent's mind is just its perception. Finally, for a B-introspective KBC system, the agent is aware of what it believes.

PROPOSITION 3. *Let $\mathcal{I}$ be an interpreted KBC system.*

If $\mathcal{I}$ is KC-introspective, then

$$\mathcal{I} \models (C\varphi \Rightarrow KC\varphi) \wedge (\neg C\varphi \Rightarrow K\neg C\varphi) \quad (1)$$

If $\mathcal{I}$ is C-introspective, then

$$\mathcal{I} \models (C\varphi \Rightarrow CC\varphi) \wedge (\neg C\varphi \Rightarrow C\neg C\varphi) \quad (2)$$

If $\mathcal{I}$ is C-introspective and CB-introspective, then

$$\mathcal{I} \models (B\varphi \Rightarrow CB\varphi) \wedge (\neg B\varphi \Rightarrow C\neg B\varphi) \quad (3)$$

If $\mathcal{I}$ is C-introspective and B-introspective, then

$$\mathcal{I} \models (B\varphi \Rightarrow BB\varphi) \wedge (\neg B\varphi \Rightarrow B\neg B\varphi) \quad (4)$$

Proof: Immediately by the satisfaction definition. ∎

3.4 Proof systems

3.4.1 The Basic KBC proof system:

The proof system contains the axioms of propositional calculus plus those formulas in Proposition 2. The proof system is closed under the propositional inference rules plus:

$$\frac{\vdash \varphi}{\vdash K\varphi \wedge C\varphi}$$

Note that we do not need to include the inference rule

$$\frac{\vdash \varphi}{\vdash B\varphi}$$

into our system because we have that $\vdash C\varphi \Rightarrow B\varphi$.

PROPOSITION 4. *The following holds in the basic KBC proof system:*

$$\vdash C\varphi \Rightarrow BK\varphi \text{ and } \vdash \neg C\varphi \Rightarrow B\neg K\varphi$$

Proof: We prove $\vdash C\varphi \Rightarrow BK\varphi$ by $\vdash C\varphi \Rightarrow CK\varphi$ and $\vdash CK\varphi \Rightarrow BK\varphi$, and get $\vdash \neg C\varphi \Rightarrow B\neg K\varphi$ by $\vdash \neg C\varphi \Rightarrow C\neg K\varphi$ and $\vdash C\neg K\varphi \Rightarrow B\neg K\varphi$. ∎

3.4.2 The introspective KBC proof system:

This proof system is the basic KBC proof system plus the formulas (1), (2) and (3) in Proposition 3.

4. COMPLETENESS AND COMPLEXITY RESULTS

In this section, we prove some fundamental results about the basic KBC proof system and the introspective KBC proof system. Because possible-worlds semantics provides a good formal tool for customizing a logic and has been well-studied for many years [20, 17, 11], we first build a bridge between interpreted KBC systems and Kripke structures. In what follows, we assume the standard definitions for Kripke structures. We refer the reader to [6, 11, 2] for a detailed exposition of the subject.

DEFINITION 5. *A Kripke structure $M = (W, \ , K, B, C)$* [1] *is called a* basic KBC Kripke structure *if*

K is an equivalence relation.

For each $w_1 \in W$, there is $w_2 \in W$ such that, for all $w \in W$, $w_1 C w$ iff $w_2 K w$.

$B \quad C$.

DEFINITION 6. *A Kripke structure $M = (W, \ , K, B, C)$ is said to be an* introspective KBC Kripke structure *if it is a basic KBC Kripke structure and satisfies the following:*

For all $w_1, w_2 \in W$, if $w_1 K w_2$, then, for all $w \in W$, $w_1 C w$ iff $w_2 K w$.

For all $w_1, w_2, w_3 \in W$, if $w_1 C w_2$, then $w_1 C w_3$ iff $w_2 C w_3$.

For each $w_1, w_2 \in W$, if $w_1 C w_2$, then for all $w \in W$, $w_1 B w$ iff $w_2 B w$.

The following lemma builds a bridge between interpreted KBC systems and Kripke structures.

LEMMA 7. *A formula φ is satisfiable by an interpreted KBC system iff it is satisfiable by a basic KBC Kripke structure. Moreover, a formula φ is satisfiable by an introspective KBC system iff it is satisfiable by an introspective KBC Kripke structure.*

[1]For convenience, we use the symbols of modalities to denote the corresponding relations in the Kripke structure

4.1 Completeness

Our first important result is the completeness of the basic KBC proof system.

THEOREM 8. *The basic KBC proof system is sound and complete with respect to interpreted KBC systems.*

Proof: The soundness part of the proof is simple and obvious; we give only the proof for the completeness part, which is inspired by the completeness proofs in [7]. We need only to prove that every KBC-consistent formula is satisfiable in a basic KBC Kripke structure.

First, we construct a special Kripke structure M^c, called *canonical* Kripke structure, as follows. Consider the set W of all maximal consistent sets of formulas. Given a $w \in W$, define

$$w/X = \{\phi \mid X\phi \in w\}$$

where X denotes one of the modalities K, B and C.

Let $M^c = (W, \ , K, B, C)$ be a Kripke structure, where

$$\begin{array}{lll} W & = & \{w : w \text{ is a maximal consistent set}\} \\ (w)(p) & = & \begin{cases} \textbf{true} & \text{if } p \in w \\ \textbf{false} & \text{if } p \notin w \end{cases} \\ X & = & \{(w, w') \mid w/X \quad w'\}, \end{array}$$

where X denotes one of K, B, and C.

We then show, by induction on the structure of ϕ, that for all w we have that

$$(M^c, w) \models \phi \text{ iff } \phi \in w. \qquad (\ )$$

More precisely, assuming that the claim holds for all subformulas of ϕ, we will show that it also holds for ϕ.

If ϕ is a primitive proposition p, this is immediate from the definition of $\ (w)(p)$ above. The cases where ϕ is a conjunction or a negation can follow easily since w is a maximal consistent set.

Then suppose that ϕ is of the form $X\varphi$ and that $\phi \in w$. So, $\varphi \in w/X$. By the definition of X in the canonical structure, if wXw', then $w/X \quad w'$ and hence $\varphi \in w'$. Thus, using the induction hypothesis, $(M^c, w') \models \varphi$ for all w' such that wXw'. So we can get by the semantical definition of X that $(M^c, w) \models \phi$.

For the other direction, if $(M^c, w) \models X\varphi$. Then the set $(w/X) \cup \{\neg\varphi\}$ must be inconsistent; otherwise, we can construct a state w' in which $(w/X) \cup \{\neg\varphi\}$ holds. By the definition of X in the canonical structure, wXw'. It follows that $(M^c, w') \models \neg\varphi$ by the induction hypothesis, so $(M^c, w) \models \neg X\varphi$ which contradicts our assumption.

Since the set of formulas $(w/X) \cup \{\neg\varphi\}$ is inconsistent, some finite subset $\{\varphi_1, \varphi_2, ..., \varphi_n, \neg\varphi\}$ of it must be inconsistent. Thus by propositional reasoning, we can have

$$\vdash \varphi_1 \Rightarrow (\varphi_2 \Rightarrow (...(\varphi_n \Rightarrow \varphi)...)).$$

By basic KBC proof rule $\frac{\vdash \varphi}{\vdash X\varphi}$, we get,

$$\vdash X(\varphi_1 \Rightarrow (\varphi_2 \Rightarrow (...(\varphi_n \Rightarrow \varphi)...))).$$

Because $\vdash X(\ ' \Rightarrow \ ) \Rightarrow (X\ ' \Rightarrow X\ )$ for any and $'$, we have, by propositional reasonings, that

$$\vdash (X\varphi_1 \Rightarrow (X\varphi_2 \Rightarrow (...(X\varphi_n \Rightarrow X\varphi)...)).$$

Thus the set $\{X\varphi_1, X\varphi_2, ..., X\varphi_n, \neg X\varphi\}$ is inconsistent. Since $\{\varphi_1, \varphi_2, ..., \varphi_n\} \in w/X$, we have $X\varphi_1, X\varphi_2, ..., X\varphi_n \in w$. Since w is a maximal consistent set, one of $X\varphi$ or $\neg X\varphi$ must be in w. Then we immediately can conclude that $X\varphi \in w$.

So for any KBC-consistent formula ϕ, it must be in some w such that $(M^c, w) \models \phi$.

We now prove that the Kripke structure M^c is a basic KBC Kripke structure. First, by the usual arguments it follows that K is an equivalence relation from the knowledge axioms and $B \subset C$ from the belief and certainty axioms. Thus, it suffices to prove that for each w_1 there is a w_2 such that for all w, $w_1 C w$ iff $w_2 K w$. Given w_1, let w_2 be a maximal consistent set that $w_1/C \subset w_2$. We need only to show that $w_1/C = w_2/K$. For each formula ϕ, if $\phi \in w_1/C$, then by the axiom $\vdash C\phi \Rightarrow CK\phi$ we have that $CK\phi \in w_1$, i.e., $K\phi \in w_1/C \quad w_2$, thus $\phi \in w_2/K$. On the other hand, if $\phi \notin w_1/C$, then $\neg C\phi \in w_1$, and by the axiom $\vdash \neg C\phi \Rightarrow C\neg K\phi$ we have that $C\neg K\phi \in w_1$, i.e. $\neg K\phi \in w_1 \quad w_2$, and hence $\phi \notin w_2/K$.

Thus, every consistent formula is satisfiable by a basic KBC Kripke structure. By Lemma 7, we have that every consistent formula is satisfiable by an interpreted KBC system. This completes the proof. ▮

As might be expected, the introspective KBC proof system characterizes completely those KBC formulas that are valid for introspective KBC systems.

THEOREM 9. *The introspective KBC proof system is sound and complete with respect to the class of introspective KBC systems.*

Proof: Similar to Theorem 8.▮

4.2 Complexity

The soundness and completeness theorems above imply that the valid problems for the both classes of interpreted KBC systems are decidable. We show that the complexity of the valid problem for general interpreted KBC systems is PSPACE-complete, but for the introspective KBC systems the complexity is much easier, indeed, it is co-NP-complete.

THEOREM 10. *The complexity of satisfiability problem for interpreted KBC systems is PSPACE-complete.*

Proof: Consider the case where K and C do not appear in the formula we want to test. Because the complexity of satisfiability problem for the modal logic K is PSPACE-complete, we immediately have that the satisfiability problem for interpreted KBC systems is PSPACE-hard. On the other hand, to prove the satisfiability problem is in PSPACE, we present a decision procedure by using Ladner's tableau method [13]. The key notion of our procedure is a *KBC tableau*, which extends the notion of *propositional tableau* [1].

A KBC tableau is a tuple $T = (S, L, \mathcal{K}, \mathcal{B}, \mathcal{C})$, where S is a set of states, $\mathcal{K}$, $\mathcal{B}$ and $\mathcal{C}$ are possibility relations, and L is a labelling function that associates with each states $s \in S$ a set $L(s)$ of formulas such that

1. $L(s)$ is a propositional tableau.
2. $K \quad \in L(s)$ and $(s,t) \in \mathcal{K}$, then $\quad \in L(t)$; and the same holds for modalities B and C.
3. $\neg K \quad \in L(s)$ then there is a t with $(s,t) \in \mathcal{K}$ and $\neg \quad \in L(t)$; and the same holds for modalities B and C.
4. (a) If $K \quad \in L(s)$, then $\quad \in L(s)$; and (b) if $(s,t) \in \mathcal{K}$, then $K \quad \in L(s)$ iff $K \quad \in L(t)$.
5. (a) If $C \quad \in L(s)$ and $(s,t) \in \mathcal{C}$, then $K \quad \in L(t)$; (b) if $\neg C \quad \in L(s)$ and $(s,t) \in \mathcal{C}$, then $\neg K \quad \in L(t)$.
6. If $C \quad \in L(s)$, then $B \quad \in L(s)$.

We say that T is a *KBC tableau for φ* if T is a KBC tableau and $\varphi \in L(s)$ for some state.

With the KBC tableaux, we can construct our decision procedure, which runs in space polynomial in the size of the input formula φ. We can trivially prove that for all formula φ with $|\varphi| = n$, the height h of the tableau tree $h \quad n^2$. Then roughly speaking, we apply the depth-first search algorithm for it and we easily get that the satisfiable tableau for φ can be computed in $O(n^3)$.▮

LEMMA 11. *A formula φ is satisfiable by an introspective KBC Kripke structure iff it is satisfiable by an introspective KBC Kripke structure with at most $2 \mid \varphi \mid$ states.*

Proof: Suppose $M = (S, \quad, K, B, C)$ be an introspective KBC Kripke structure, s_0 is a state of M with $(M, s_0) \models \varphi$. For each X of K, B and C, let F_X be the set of sub-formulas of φ of the form $\quad$ for which $(M, s_0) \models \neg X \quad$. For each $\quad \in F_X$, there must be a state $s_X^\psi \in S$ such that $(M, s_X^\psi) \models \neg \quad$. Let $S_X = \{s_X^\psi \mid \quad \in F_X\}$. Let $M' = (S', \quad', K', B', C')$, where $S' = \{s_0\} \cup S_K \cup S_C$, $\quad'$ is the restriction of $\quad$ to S', $K' = \{(s,t) \mid s,t \in \{s_0\} \cup S_K \text{ or } s,t \in S_C\}$, $B' = \{(s,t) \mid s \in S', t \in S_B\}$, and $C' = \{(s,t) \mid s \in S', t \in S_C\}$. Since $\mid F_K \mid < |\varphi|$ and $\mid F_C \mid < |\varphi|$, it follows that $S' \quad 2|\varphi|$.

We first show that the following claims hold.

1. $S_B \quad S_C$
2. (a) For all $s, t \in \{s_0\} \cup S_K$, $(s,t) \in K$; (b) For all $s, t \in S_C$, $(s,t) \in K$.
3. For all $s \in S'$ and for all $t \in S_B$, $(s,t) \in B$.
4. For all $s \in S'$ and for all $t \in S_C$, $(s,t) \in C$.

The first claim is by the property $B \quad C$ of basic KBC Kripke structures. Part (a) of the second claim follows by the construction S_K and that K is an equivalence relation. Part (b) is because M is also by the definition of basic KBC Kripke structures and $(s_0, s), (s_0, t) \in C$. The last two hold for that M is an introspective Kripke structure.

We now show that for all states $s' \in S'$ and for all sub-formulas $\quad$ of φ, $(M, s') \models \quad$ iff $(M', s') \models \quad$. We proceed by induction on the structure of $\quad$. The nontrivial cases are when $\quad$ is of the form $K \quad'$, $B \quad'$ or $C \quad'$. Suppose $s' \in S'$.

> If $(M, s') \models K \quad'$, then, $(M, t) \models \quad'$ for all $t \in S$ with $(s', t) \in K$. Thus, by the induction hypothesis, $(M', t) \models \quad'$ for all $t \in S$ with $(s', t) \in K$. Hence, by the second claim, $(M', t) \models \quad'$ for all $t \in S'$ with $(s', t) \in K'$, that is, $(M', s') \models K \quad'$. And if $(M, s') \not\models K \quad'$, then $(M, s') \models \neg K \quad'$. There are two cases: $s' \in \{s_0\} \cup S_K$ or $s' \in S_C$. If $s' \in \{s_0\} \cup S_K$, then $(s_0, s') \in K$, and hence $(M, s_0) \models \neg K \quad'$. By the construction, $s_K^{\psi'} \in S_K$ and $(M, s_K^{\psi'}) \not\models \quad'$. By the induction hypothesis, $(M', s_K^{\psi'}) \not\models \quad'$, and hence $(M', s') \not\models K \quad'$. On the other hand, if $s' \in S_C$, then $(M, s_0) \models C\neg K \quad'$. Because M is a basic KBC Kripke structure, we have $(M, s_0) \models \neg C \quad'$. Thus, $s_C^{\psi'} \in S_C$ and $(M, s_C^{\psi'}) \models \neg \quad'$. By the induction hypothesis, we have $(M', s_C^{\psi'}) \models \neg \quad'$. By the second claim again, we have that $(M', s') \not\models K \quad'$.
>
> If $(M, s') \models C \quad'$, then, $(M, t) \models \quad'$ for all $t \in S$ with $(s', t) \in C$. Thus, by the induction hypothesis, $(M', t) \models \quad'$ for all $t \in S$ with $(s', t) \in C$. Hence, by the fourth claim, we have $(M', t) \models \quad'$ for all $t \in S'$ with $(s', t) \in C'$, that

is $(M', s') \models C\ \ '$. And if $(M, s') \not\models C\ \ '$, then $(M, s') \models \neg C\ \ '$, because either $(s_0, s') \in K$ or $(s_0, s' \in C)$, we have that $(M, s_0) \models \neg C\ \ '$. Thus, $s_C^{\psi'} \in S_C$ and $(M, s_C^{\psi'}) \not\models\ \ '$. By the induction hypothesis, $(M', s_C^{\psi'}) \not\models\ \ '$. Hence, $(M', s') \not\models C\ \ '$.

By the same argument as above, we get that $(M, s') \models B\ \ '$ iff $(M', s') \models B\ \ '$. ∎

THEOREM 12. *The complexity of satisfiability problem for introspective KBC systems is NP-complete.*

Proof: Cook's theorem [4] implies that the complexity of satisfiability problem for introspective KBC systems is NP-hard. By Lemma 11, it is easy to construct a nondeterministic polynomial algorithm for deciding whether a formula φ is satisfiable. ∎

5. RELATED WORK AND CONCLUSIONS

5.1 Logics of knowledge, belief and certainty

The notion of certainty used in this paper has been first introduced by Lamarre and Shoham [14] and similar notions are Lenzen's *strong belief* [15] and Voorbraak's *rational belief* [21]. Lenzen [15] lists many of the syntactic properties of the notions of knowledge, belief and certainty (i.e., strong belief), but it does not provide any semantics. Lamarre and Shoham [14] provide a model theory of knowledge, belief and certainty, with respect to which all Lenzen's collection of axioms are valid; however they reject S5 as the logic describing the knowledge operator and their logic is not computationally grounded.

5.2 Computationally grounded logics

Perhaps closest to our work in this paper are some computationally grounded logics in the field of agent theory. In the mid 1980s, Halpern and his colleagues discovered that S5 epistemic logics could be given a natural interpretation via interpreted systems model [6, 9, 8]. Interpreted systems are very close to our interpreted KBC systems; interpreted systems play the role in epistemic logic just as interpreted KBC systems do in KBC logic. However, our construction contains more elements than theirs. We distinguish between what is visible of the environment and what an agent is seeing or perceiving. Also, we distinguish between what the agent perceives about the visible part of the environment and what, the agent conjectures, the invisible part of the environment could be like.

Another computationally grounded logic is $\mathcal{VSK}$ logic [23], which enables us to represent what is *visible* of the environment to individual agents, what these agents actually *perceive* (see), and what the agents actually *know* about the environment. Wooldridge and Lomuscio's visibility operator $\mathcal{V}$ in $\mathcal{VSK}$ logic corresponds to our knowledge operator K, and the perception operator $\mathcal{S}$ to our certainty operator C. The key differences are that they adopt S5 system for the perception operator, while we adopt only the K (in the basic KBC proof system) or K45 (in the introspective proof system) system for the certainty operator. Intuitively, their interpretation of $\mathcal{S}\varphi$ is that *the perception received by the agent carries the information* φ; while $C\varphi$ in this paper means that, according to the received perception, the agent feels certain of φ. Thus, Wooldridge and Lomuscio's notion of perception is an external one, while ours is internal. In addition, they did not consider the notion of belief, which we think is useful to formalize other notions of agent's mental state such as that of goal. Moreover, all modalities in $\mathcal{VSK}$ logic satisfy the S5 axioms. Thus, it is interesting to investigate those computationally grounded modal logics that not only formalize external notions via S5 axioms but also characterize an the internal attitudes of an agent–its beliefs, desires, etc, beyond S5 axioms.

The deontic interpreted system model [16] can also be regarded as computationally grounded one, which is closely related to the agent-environment system [23] and the interpreted KBC system in this paper, since the so-called red states may be thought of as those states where the agent get the wrong perception and hence may behave incorrectly.

5.3 Complexity of modal logics

The worst-case complexity of modal logics is a flouring research activity, and it is impossible for us to list the literature here. A good overview can be found in [19]. It is well-known that S5 or K45 logic is NP-complete (no harder than that of the propositional logic) for their satisfiability problem [13] and the complexity of the satisfiability problem for multimodal logic with two or more independent S5 or K45 operators becomes PSPACE-complete [7]. Thus, it is very interesting to build those multimodal logics with NP-complete satisfiability problem. In [5], it is proved that if the equivalence relations are *ordered locally*, then the multimodal logic with n S5 modalities preserves NP-completeness of satisfiability problem. This paper presents a new multimodal logic for which the satisfiability problem is NP-complete.

5.4 Concluding remarks

We have developed a logic of knowledge, belief and certainty, which allows us to explicitly mention an agent's knowledge, belief and certainty in multi-agent systems. A computationally grounded model, called interpreted KBC systems, is given for interpreting this logic. To characterize all valid formulas in our logic, we have provided a sound and complete proof system. We also have presented a procedure deciding whether a formula is valid with the computational complexity of PSPACE, by using the well-known tableau method. We identify a subclass of interpreted KBC systems, called *introspective KBC systems*. The validity problem for introspective KBC systems turns out to be co-NP complete, and is no harder than that of the propositional logic. We have also given a sound and complete proof system with respect to introspective KBC systems.

We are currently working on the analysis and verification of security protocols by using the multiagent version of the KBC logic. Our future work also includes the temporal extension to the logic, the formalizations of the other notions of an agent's mental state such as 'goal', and proof theories for other classes of interpreted KBC systems (for example, the class of C-introspective and CB-introspective systems).

Acknoledgement

We would like to thank the anonymous referees for insightful and helpful comments. Special thanks to Ruqian Lu whose suggestions inspire us to write an earlier version of this paper.

6. REFERENCES

[1] E. Beth. *The Foundation of Mathematics*. North-Holland, Amsterdam, 1959.

[2] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.

[3] P. Cohen and H. Levesque. Intension is choice with commitment. *Artificial Intelligence*, 42:23–261, 1990.

[4] S. Cook. The complexity of theorem proving procedure. In *Proc. 3rd ACM Symp. on Theory of Computing*, pages 151–158, 1971.

[5] S. Demri. A class of decidable information logics. *Theoretical Computer Science*, 195(1):33–60, 1998.

[6] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning about knowledge*. MIT Press, Cambridge, MA, 1995.

[7] J. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54:319–379, 1992.

[8] J. Halpern and M. Vardi. The complexity of reasoning about knowledge and time: extended abstract. In *Proc. 18th Annual ACM Symposium on Theory of Computing*, pages 304–315, 1986.

[9] J. Halpern and L. Zuck. A little knowledge goes a long way: Simple knowledge based derivations and correctness proofs for a family of protocols. *J. ACM*, 39(3):449–478, 1992.

[10] J. Hintikka. *Knowledge and Belief*. Cornell University Press, Ithaca, NY, 1962.

[11] G. Hughes and M. Cresswell. *A New Introduction to Modal Logic*. London: Routledge., 1996.

[12] S. Kripke. A semantical analysis of modal logic. i: Normal modal propositional calculi. *Z. Math. Logik Grundl. Math.*, 9:67–96, 1963.

[13] R. Ladner. The computational complexity of provability in systems of modal propositional logic. *SIAM Journal on Computing*, 6(3):467–480, 1977.

[14] P. Lamarre and Y. Shoham. Knowledge, certainty, belief, and conditionalisation. In *KR94*, pages 415–424. Morgan Kaufmann, 1994.

[15] W. Lenzen. *Recent Work in Epistemic Logic*, volume XXX of *Acta Philosophica Fennica*. North-Holland, Amsterdam, 1978.

[16] A. Lomuscio and M. Sergot. Deontic interpreted systems. *Studia Logica*, 75, 2003.

[17] J.-J. Meyer and W. van der Hoek. *Epistemic Logic for Computer Science and Artificial Intelligence*. Cambridge Tracts in Theoretical Computer Science 41. Cambridge University Press, 1995.

[18] A. Rao and M. Georgeff. Decision procedures for BDI logics. *Journal of Logic and Computation*, 8(3):293–344, 1998.

[19] E. Spaan. *Complexity of Modal Logics*. PhD thesis, Amsterdam University, ILLC, March 1993.

[20] J. van Benthem. Correspondence theory. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume II, pages 167–247. D. Reidel Publ. Co., 1984.

[21] F. Voorbraak. The theory of objective knowledge and rational belief. In *Logics in AI, European Workshop JELIA '90*, pages 499–515. Springer Verlag, 1991.

[22] M. Wooldridge. Computationally grounded theories of agency. In E. Durfee, editor, *Proceedings of the Fourth International Conference on Multi-Agent Systems (ICMAS 2000)*. IEEE Press, 2000.

[23] M. Wooldridge and A. Lomuscio. A computationally grounded logic of visibility, perception, and knowledge. *Logic Journal of the IGPL*, 9(2):273–288, 2001.