

# An Extended Interpreted System Model for Epistemic Logics

Kaile Su<sup>1,2</sup> and Abdul Sattar<sup>2</sup>

<sup>1</sup>Key laboratory of High Confidence Software Technologies (Peking University), Ministry of Education, Beijing, China

<sup>2</sup>Institute for Integrated and Intelligent Systems, Griffith University, Brisbane, Australia  
sukl@pku.edu.cn; a.sattar@griffith.edu.au

## Abstract

The interpreted system model offers a computationally grounded model, in terms of the states of computer processes, to S5 epistemic logics. This paper extends the interpreted system model, and provides a computationally grounded one, called the interpreted perception system model, to those epistemic logics other than S5. It is usually assumed, in the interpreted system model, that those parts of the environment that are visible to an agent are correctly perceived by the agent as a whole. The essential idea of the interpreted perception system model is that an agent may have incorrect perception or observations to the visible parts of the environment and the agent may not be aware of this. The notion of knowledge can be defined so that an agent knows a statement iff the statement holds in those states that the agent can not distinguish (from the current state) by using only her correct observations. We establish a logic of knowledge and certainty, called KC logic, with a sound and complete proof system. The knowledge modality in this logic is S4 valid. It becomes S5 if we assume an agent always has correct observations; and more interestingly, it can be S4.2 or S4.3 under other natural constraints on agents and their sensors to the environment.

## Introduction

Epistemic logics or logics of knowledge have been studied by philosophers for a long time, and with the advent of agent oriented computing, they are of paramount importance for the formalization of autonomous agents. There are two main semantic approaches to epistemic logics, the *possible worlds semantics* (Hintikka 1962; Kripke 1963) and the *interpreted system model* (Fagin et al. 1995; Halpern and Zuck 1992; Halpern and Vardi 1986). The first approach is very fruitful, including S4, S4.2, S4.3 etc. The advantage of this approach is that various properties about the notion of knowledge can be characterized conveniently with a model theoretic feature in terms of the accessibility relations. The second approach, mainly due to Halpern and his colleagues (Fagin et al. 1995; Halpern and Zuck 1992; Halpern and Vardi 1986), offers a very natural interpretation, in terms of the states of computer processes, to S5 epistemic logics. The salient point of the second approach is that we are able to associate the sys-

tem with a computer program and formulas can be understood as properties of program computations. In this sense, the interpreted system model is *computationally grounded* (Wooldridge 2000).

However, the epistemic logic characterized by the interpreted system model is forced to be S5, as the undistinguishability relation between two global states in an interpreted system is an equivalence one. A natural question is whether the interpreted system model can be extended as a natural and satisfying semantic model for those epistemic logics other than S5 such as S4, S4.2 and S4.3.

This paper extends the interpreted system model, and provides a computationally grounded one, called the *interpreted perception system model*, to those epistemic logics other than S5. The essential idea of the interpreted perception system model is that an agent may have incorrect observations to the environment and the agent may not be aware of this. On the contrary, in the interpreted system model, it is usually assumed that the visible parts of the environment are correctly perceived as a whole; as a result, the visible parts of the environment and the perceptions of the agent may be modelled as shared variables in both the design module of the environment and that of the agent. Our notion of knowledge can be defined so that an agent knows a statement iff the statement holds in those states that the agent can not distinguish (from the current state) by using only her correct observations. The resulting logic for the knowledge modality is S4. It becomes S5 if we assume an agent always has only correct observations to the environment; and more interestingly, it can be S4.2 or S4.3 under other natural constraints on agents and their observations.

Our philosophical hypothesis about knowledge is that an agent knows a statement iff the statement is true, the agent is certain of the statement and the agent is certain of the statement for good reasons. An important question is then raised here, “what are good reasons or bad reasons?” In our concrete computational model, we can naturally take the so-called “good reasons” to be correct observations. As a result, an agent’s knowledge is directly from her correct observations or is derived by her correct observations with her knowledge about the system.

The significance of this work is in three aspects. First, we provide a computational approach to evaluating the statements in epistemic logic for an agent with possibly faulty

sensors. Secondly, we explore some particular classes of the interpreted perception systems, which are exactly related to several most influential epistemic logics in the philosophic community. Finally, although the notion of knowledge in Artificial Intelligence (AI) is generally (though not universally) S5 valid (Lamarre and Shoham 1994; Meyer and van der Hoek 1995), we show, to some content, that epistemic logics other than S5 can also be applicable in AI, because we endow them with a concrete computational model.

To illustrate the importance of a concrete computational model for the notion of knowledge, let us consider the scenario where an Intelligent Electronic Tutor (IET) values Alice's knowledge about " $2 + 2 = 4$  or  $2 + 3 \neq 5$ " after displaying both statements " $2 + 2 = 4$ " and " $2 + 3 = 5$ ". IET may ask Alice the following questions:

Are you certain of " $2 + 2 = 4$  or  $2 + 3 \neq 5$ "? why?

If Alice's answer is "Yes, because I observed " $2 + 2 = 4$ " and " $2 + 3 \neq 5$ "", then IET will think that Alice knows " $2 + 2 = 4$  or  $2 + 3 \neq 5$ ". Here Alice has two reasons for her certainty about " $2 + 2 = 4$  or  $2 + 3 \neq 5$ ". Though one reason is incorrect, the other is enough to deduce the statement " $2 + 2 = 4$  or  $2 + 3 \neq 5$ ". If Alice's answer is "Yes, because I observed " $2 + 2 \neq 4$ " and " $2 + 3 \neq 5$ ", then IET will not think that Alice knows " $2 + 2 = 4$  or  $2 + 3 \neq 5$ ", though this statement is true and she is certain of it.

To make IET genuinely intelligent, a concrete computational approach is essential for systemically evaluating a student's knowledge. There are other more serious cases (legal reasoning, for instance) that call for a computational approach to evaluate the statements in epistemic logic like "if an agent *knows* that a document is secret, then the agent will not reveal the document".

In this paper, we deliberately consider only one agent and only static epistemic properties. However, to demonstrate the intuition behind the notion of knowledge, our epistemic logic includes another modality for the notion of *certainty*, which is a kind of belief and it might be false, but an agent who feels certain of a proposition is not aware that the proposition might be false.

The remainder of this paper is organized as follows. In the next section, we extend the interpreted system model and get the so-called the interpreted perception system model. Then, a sound and complete axiom system is established with respect to this semantic model. We also define several specific classes of uncertain systems, and relate them to the standard epistemic logics S4.2 and S4.3. Finally, we discuss the related work and conclude the paper.

## The Interpreted Perception System Model

Let us first briefly review the notion of *Interpreted System*, which is originally defined by Halpern and Moses (Fagin et al. 1995). The fundamental notion on which interpreted systems are defined is the one of 'local state'. Intuitively, the local state of an agent represents the entire information about the system observed by the agent. The (global) state of the system is defined as the pair of the local states of the agent and that for the environment.

Formally, a *system of global states* is a non-empty subset of Cartesian  $L_e \times L_a$ , where  $L_e$  represents a set of local states for the environment and  $L_a$  is a set of the agent's local states.

Assume that we have a set  $\Phi$  of primitive propositions, which we can think of as describing basic facts about the system. An *interpreted system of global states* consists of a pair  $(S, \pi)$ , where  $S$  is a system of global states and  $\pi$  is a valuation function, which assigns truth values to primitive propositions.

The notion of an *interpreted system* in the framework in (Fagin et al. 1995) represents the temporal evolution of a system by means of *runs*. A run, in their terminology, is a function from the set of natural numbers to the set global states, and an interpreted system is a set of runs over global states together with a valuation for the primitive propositions. In this paper, we do not deal with time, and work only on states for simplicity.

We now define *interpreted perception systems* by extending the notion of interpreted systems. The basic idea is that an agent may have incorrect observations to her environment. More specifically, we assume that the agent has  $n$  sensors, and divide the visible part of the environment into  $n$ -tuple so that the  $i$ -th ( $i < n$ ) sensor perceives the  $i$ -th element of the visible part of the environment. As the agent may have faulty sensors, we distinguish what the agent sees or perceives from the visible part of the environment. For convenience, we call the visible part of the environment the *accessible state* of the agent through out this paper.

Given an accessible state  $(l_0, \dots, l_{n-1})$ , and the agent's perception  $(l'_0, \dots, l'_{n-1})$ , if for some  $i < n$ ,  $l_i \neq l'_i$ , then the agent's  $i$ -th sensor should be faulty.

Formally, we have the following definitions.

**Definition 1** An *extended global state* is a triple  $(l_e, l_a, l'_a)$ , where  $l_e$  is a local state for the environment,  $l_a$  a  $n$ -tuple representing an accessible state of the agent, and  $l'_a$  a  $n$ -tuple representing the agent's perception to her accessible state.

Given an extended global state  $g = (l_e, l_a, l'_a)$ , we use  $Env(g)$ ,  $Acc(g)$ , and  $Per(g)$  to denote  $l_e$ ,  $l_a$ , and  $l'_a$ , respectively. For every  $i < n$ , we denote the  $i$ -th elements of  $l_a$  and  $l'_a$  by  $Acc_i(g)$  and  $Per_i(g)$ , respectively.

Intuitively, if  $Acc_i(g) = Per_i(g)$ , then the  $i$ -sensor gets the correct perception for the  $i$ -th attribute value of the environment. Let  $\mathbb{C}_g$  be the set of those  $i < n$  such that  $Acc_i(g) = Per_i(g)$ .

For convenience, sometimes we still use "global states" for "extended global states" in this paper.

**Definition 2** A perception system is a non-empty set of extended global states. An *interpreted perception system* consists of a pair  $(S, \pi)$ , where  $S$  is a perception system and  $\pi$  is a valuation function such that, for every  $g \in S$  and  $p \in \Phi$ ,  $\pi(Env(g), Acc(g))(p) \in \{\text{true}, \text{false}\}$ .

## A Logic of Knowledge and Certainty

In this section, we introduce a modal logic of knowledge and certainty, called *KC* logic. The semantics for *KC* logic is given by using the perception system model. In this semantics, we capture the notion of knowledge by distinguishing

correct perceptions from incorrect ones, so that the agent knows  $\varphi$  iff the agent can judge  $\varphi$  from the correct part of her perceptions.

### Syntax

Given a set  $\Phi$  of propositional atoms, the language of  $KC$  logic is defined by the following BNF grammar:

$$\langle wff \rangle ::= \text{any element of } \Phi \mid \neg \langle wff \rangle \mid \langle wff \rangle \wedge \langle wff \rangle \mid K \langle wff \rangle \mid C \langle wff \rangle$$

Formula  $K\phi$  means that the agent knows  $\phi$  or can judge  $\varphi$  from the correct part of her perceptions about the environment.  $C\phi$  means that the agent feels for certain the property  $\phi$ . Intuitive meaning behind ‘certainty’ is that, to the agent, the facts of which he is certain appear to be knowledge (Lamarre and Shoham 1994).

### Semantics

We now propose a kind of semantics to interpret the  $KC$  logic formulas in terms of interpreted perception systems. In the following, we define the satisfaction relation  $\models_{KC}$  between a formula and a pair of an interpreted perception system and an extended global state. Given an interpreted system  $\mathcal{I} = (S, \pi)$ , and an extended global state  $g \in S$ , we have that:

- $(\mathcal{I}, g) \models_{KC} p$  iff  $\pi(Env(g), Acc(g))(p) = \mathbf{true}$ , where  $p$  is a primitive proposition in  $\Phi$ ;
- $(\mathcal{I}, g) \models_{KC} \neg\varphi$  iff  $(\mathcal{I}, g) \not\models_{KC} \varphi$ ;
- $(\mathcal{I}, g) \models_{KC} \varphi \wedge \psi$  iff  $(\mathcal{I}, g) \models_{KC} \varphi$  and  $(\mathcal{I}, g) \models_{KC} \psi$ ;
- $(\mathcal{I}, g) \models_{KC} K\varphi$  iff  $(\mathcal{I}, g') \models_{KC} \varphi$  for those  $g' \in S$  such that  $Per(g') = Per(g)$ , and  $Acc_i(g') = Acc_i(g)$  for all  $i \in \mathbb{C}_g$ .
- $(\mathcal{I}, g) \models_{KC} C\varphi$  iff  $(\mathcal{I}, g') \models_{KC} \varphi$  for those  $g' \in S$  such that  $Acc(g') = Per(g)$  and  $Per(g') = Per(g)$ .

Informally speaking, an agent knows  $\varphi$  iff  $\varphi$  holds at all those global states where the agent has the same perception as at the current global state, and any visible environment attribute value is either the same as in the current or perceived incorrectly by the agent at the current state. This means that an agent deduces her knowledge by her correct perceptions and the system properties that are known to her.

As for the intuition behind modality  $C$ , we remark that  $C\varphi$  iff  $\varphi$  holds at those global states where the agent has the same perception as at the current global state, and all attribute values for the visible part of the environment is perceived correctly. Thus, an agent feels certain of a statement if the statement can be deduced by her perceptions and the system properties that are known to her under the assumption that her perceptions are all correct.

Given two global states  $g$  and  $g'$ , we call  $g'$  *g-undistinguished*, if  $Per(g') = Per(g)$  and  $Acc_i(g') = Acc_i(g)$  for all  $i \in \mathbb{C}_g$ . Therefore, an agent knows  $\varphi$  at the current global state  $g$  iff  $\varphi$  holds at all  $g$ -undistinguished global states in the underlying system.

We now check what are the valid formulas for semantics  $\models_{KC}$ .

**Proposition 3** *The following formulas that are valid with respect to  $\models_{KC}$ :*

- *knowledge*  
 $K(\varphi \Rightarrow \psi) \Rightarrow (K\varphi \Rightarrow K\psi)$   
 $K\varphi \Rightarrow \varphi$   
 $K\varphi \Rightarrow KK\varphi$
- *certainty*  
 $C(\varphi \Rightarrow \psi) \Rightarrow (C\varphi \Rightarrow C\psi)$
- *knowledge and certainty*  
 $K\varphi \Rightarrow C\varphi$   
 $(C\varphi \Rightarrow CK\varphi) \wedge (\neg C\varphi \Rightarrow C\neg K\varphi)$   
 $(C\varphi \Rightarrow KC\varphi) \wedge (\neg C\varphi \Rightarrow K\neg C\varphi)$

From the proposition above, we can see that knowledge entails certainty and knowledge satisfies S4 properties for semantics  $\models_{KC}$ . The following example shows that S5 does not hold for knowledge.

**Example 4** Let us consider the scenario in the introduction section again. We assume Alice has two basic observations. The first observation is for getting the truth value of statement “ $2 + 2 = 4$ ”, and the other for getting that of statement “ $2 + 3 = 5$ ”. We use  $l_1^+$  and  $l_2^+$  to denote that statements “ $2 + 2 = 4$ ” and “ $2 + 3 = 5$ ” hold, respectively. Similarly,  $l_1^-$  and  $l_2^-$  denote the falseness of statement “ $2 + 2 = 4$ ” and that of “ $2 + 3 = 5$ ”, respectively. Finally, we assume that the environment has a unique local state, denote by  $e$ .

We now define a perception system  $S_0$  as follows:

$$S_0 = \{(e, (l_1, l_2), (l'_1, l'_2)) \mid l_i, l'_i \in \{l_i^+, l_i^-\} \text{ for } i = 1, 2\}.$$

Let  $p_1$  and  $p_2$  be the primitive propositions, which represent statements ‘ $2 + 2 = 4$ ’ and ‘ $2 + 3 = 5$ ’, respectively. The valuation function  $\pi$  is such that, for every extended global state  $g = (e, (l_1, l_2), (l'_1, l'_2))$ , and  $i = 1, 2$ ,  $\pi(e, (l_1, l_2))(p_i) = \mathbf{true}$  iff  $l_i = l_i^+$ . Thus, we get a perception system system  $\mathcal{I}_0 = (S_0, \pi)$ .

Let us consider the global state  $g_0 = (e, (l_1^+, l_2^+), (l_1^+, l_2^-))$ , which indicates Alice is certain of “ $2 + 2 = 4$ ” and “ $2 \times 2 \neq 4$ ”. It is easy to check that

- $\mathcal{I}_0, g_0 \models_{KC} K(p_1 \vee p_2)$ , and
- $\mathcal{I}_0, g_0 \models_{KC} \neg Kp_2 \wedge \neg K\neg Kp_2$ .

The latter indicates that the negative introspection axiom of S5 does not hold in the interpreted perception system model.

### Proof System

We now propose a proof system, denoted by  $\vdash_{KC}$ , for characterizing valid formulas with respect to semantics  $\models_{KC}$ . The proof system contains the axioms of propositional calculus plus those formulas in Proposition 3. The proof system is closed under the propositional inference rules plus:

$$\frac{\vdash_{KC} \varphi}{\vdash_{KC} K\varphi}.$$

Note that the inference rules for modality  $C$  is unnecessary because we have  $\vdash_{KC} K\varphi \Rightarrow C\varphi$ . Here are some interesting consequences of the proof system.

**Proposition 5** *The following hold proof system  $\vdash_{KC}$ :*

1.  $\vdash_{KC} \neg C \text{false} \Rightarrow (C\varphi \Leftrightarrow \neg K\neg K\varphi)$
2.  $\vdash_{KC} C\varphi \Leftrightarrow CK\varphi$

**Proof:** By  $\vdash_{KC} K\varphi \Rightarrow C\varphi$ , we obtain that  $\vdash_{KC} \neg K\neg K\varphi \Rightarrow \neg K\neg C\varphi$ . However, we have that  $\vdash_{KC} K\neg C\varphi \Rightarrow \neg C\varphi$  and  $\vdash_{KC} \neg C\varphi \Rightarrow K\neg C\varphi$ . It follows that  $\vdash_{KC} K\neg C\varphi \Leftrightarrow \neg C\varphi$ . Thus,  $\vdash_{KC} \neg K\neg K\varphi \Rightarrow C\varphi$ . On the other hand,  $\vdash_{KC} C\varphi \Rightarrow CK\varphi$ ; and  $\vdash_{KC} K\neg K\varphi \Rightarrow C\neg K\varphi$ . Thus,  $\vdash_{KC} C\varphi \wedge K\neg K\varphi \Rightarrow C(K\varphi \wedge \neg K\varphi)$ . It follows that  $\vdash_{KC} \neg C \text{false} \Rightarrow (C\varphi \Rightarrow \neg K\neg K\varphi)$ . Therefore, we obtain the first part of the proposition.

The second part follows from  $\vdash_{KC} C\varphi \Rightarrow CK\varphi$  and  $\vdash_{KC} K\varphi \Rightarrow \varphi$ . ■

## Completeness Result

In this section, we prove the completeness result about proof system  $\vdash_{KC}$ . Because possible-worlds semantics provides a good formal tool for customizing a logic and has been well-studied for many years (van Benthem 1984; Meyer and van der Hoek 1995; Hughes and Cresswell 1996), we first build a bridge between interpreted perception systems and Kripke structures.

### KC Kripke Structures

We now consider those Kripke structures corresponding to interpreted perception systems. We assume the standard definitions for Kripke structures. We refer the reader to (Fagin et al. 1995; Hughes and Cresswell 1996; Blackburn, de Rijke, and Venema 2001) for a detailed exposition of the subject.

**Definition 6** A Kripke structure  $M = (W, \pi, K, C)$ <sup>1</sup> is called a *KC Kripke structure* if

- $K$  is reflexive and transitive relation.
- $C \subseteq K$ .
- For all  $w_1, w_2, w_3 \in W$ , if  $w_1 C w_2$ , then  $w_1 C w_3$  iff  $w_2 K w_3$ .
- For all  $w_1, w_2, w_3 \in W$ , if  $w_1 K w_2$ , then,  $w_1 C w_3$  iff  $w_2 C w_3$ .

The following lemma builds a bridge between interpreted perception systems and *KC* Kripke structures.

**Lemma 7** A formula  $\varphi$  is satisfiable with respect to  $\models_{KC}$  iff it is satisfiable by a *KC* Kripke structure.

**Proof:** We first show that a formula  $\varphi$  is satisfiable in an interpreted KC system  $\mathcal{I}$  with respect to  $\models_{KC}$ , then it is satisfiable by a *KC* Kripke structure. Given an interpreted perception system  $\mathcal{I} = (\mathcal{R}, \pi)$ . We define a *KC* Kripke structure  $M_{\mathcal{I}} = (W, K, C, \pi')$  as follows:

1.  $W$  is the set of extended global states of  $\mathcal{I}$ .
2. For all  $w_1, w_2 \in W$ , we identify relations  $K$  and  $C$  as follows:
  - (a)  $w_1 K w_2$  iff  $Per(w_2) = Per(w_1)$ , and  $Acc_i(w_2) = Acc_i(w_1)$  for all  $i \in \mathbb{C}_{w_1}$ ; and

<sup>1</sup>For convenience, we use the symbols of modalities to denote the corresponding relations in the Kripke structure.

- (b)  $w_1 C w_2$  iff  $Acc(w_2) = Per(w_1)$  and  $Per(w_2) = Per(w_1)$ .

3.  $\pi'$  is such that, for all  $w \in W$  and primitive proposition  $p$ ,

$$\pi'(w)(p) = \pi(Acc(w), Env(w))(p).$$

For every formula  $\varphi$ , we can prove, by induction on  $\varphi$ ,  $(\mathcal{I}, g) \models_{KC} \varphi$  iff  $(M_{\mathcal{I}}, g) \models \varphi$ . Thus, if a formula is satisfiable by an interpreted perception system with respect to semantics  $\models_{KC}$ , then it is satisfiable by a *KC* Kripke structure.

Now we show that if a formula is satisfiable by a *KC* Kripke structure  $\varphi$  then it is satisfiable by an interpreted perception system with respect to semantics  $\models_{KC}$ . Given a *KC* Kripke structure  $(W, K, C, \pi')$ . We construct a corresponding interpreted perception system  $\mathcal{I}_M = (S, \pi)$  as follows.

First, for every possible world  $w \in W$ , we set a sensor for the agent, which tells the agent whether the current world is distinguishable from  $w$ . For convenience, we enumerate the  $|W|$  many sensors, and for each  $w \in W$ , we have a number  $i_w < |W|$  and let the  $i_w$ -th sensor of the agent be related to the possible world  $w$ .

For every possible world  $w \in W$ , we define an extended global state  $g_w = (l_e, l_a, l'_a)$  such that

1.  $l_e = w$ .
2. For each  $w' \in W$ , the  $i_{w'}$ -th element of  $l_a$  is 1 if  $w K w'$ , otherwise, 0.
3. For each  $w' \in W$ , the  $i_{w'}$ -th element of  $l'_a$  is 1 if  $w C w'$ , otherwise, 0.

Let the set  $S$  of extended global states be  $S = \{g_w \mid w \in W\}$ . The valuation function  $\pi$  is such that, for every  $g_w \in S$  and primitive proposition  $p$ ,  $\pi(Env(w), Acc(w))(p) = \pi'(w)(p)$ . Finally, we get the interpreted perception system  $\mathcal{I}_M = (S, \pi)$ .

Again, for every formula  $\varphi$ , we can prove, by induction on  $\varphi$ ,  $(\mathcal{I}_M, g_w) \models \varphi$  iff  $(M, w) \models \varphi$ . This completes the outline of the proof. ■

### The Completeness Proof

We now present an important technical result, the soundness and completeness of the *KC* proof system.

**Theorem 8** The *KC* proof system is sound and complete with respect to interpreted perception 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 (Halpern and Moses 1992). By Lemma 7, we need only to prove that every formula consistent with the *KC* proof system is satisfiable in a *KC* 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  $K$  and  $C$ .

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

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

where  $X$  denotes one of  $K$  and  $C$ .

We can show as in the completeness proof for standard normal modal logics, by induction on the structure of  $\phi$ , that for all  $w$  we have that

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

Finally, we can prove that the Kripke structure  $M^c$  is a  $KC$  Kripke structure. This completes the outline of the proof. ■

## Several Classes of Perception Systems

The notion of knowledge captured in the interpreted perception model satisfies S4 settled by (Hintikka 1962). Moreover, for those agents without any faulty sensor, the interpreted perception system model degenerates into the interpreted system model, and the knowledge modality is S5 valid. In this section, we present several specific classes of perception systems, which are related to some influential epistemic logics between S4 and S5.

**Definition 9** Let  $S$  be a perception system.

1. We say that  $S$  is *directed*, if for all  $g$ -undistinguished global states  $g_1$  and  $g_2$  in  $S$ , there is a global state  $g_3$  in  $S$  that is both  $g_1$ -undistinguished and  $g_2$ -undistinguished.
2.  $S$  is called *connected*, if for all  $g$ -undistinguished global states  $g_1$  and  $g_2$  in  $S$ , we have that  $\mathbb{C}_{g_1} \subseteq \mathbb{C}_{g_2}$  or  $\mathbb{C}_{g_2} \subseteq \mathbb{C}_{g_1}$ .

Clearly,  $S$  is directed if connected. Moreover, in the case where  $S$  is finite, the directedness of  $S$  means that for every global state  $g$ , there is a  $g$ -undistinguished global state where the agent has most correct observations.

Let  $\vdash_{KC+d}$  and  $\vdash_{KC+c}$  denote the proof systems resulting from  $\vdash_{KC}$  by adding formulas  $\neg K \neg K \varphi \Rightarrow K \neg K \neg \varphi$  and  $K(K\varphi_1 \Rightarrow K\varphi_2) \vee K(K\varphi_2 \Rightarrow K\varphi_1)$ , respectively. It can be proved that Proof system  $\vdash_{KC+d}$  ( $\vdash_{KC+c}$ ) is sound and complete with respect to those interpreted perception systems where the underlying perception systems are directed (connected).

Let us consider one more specific class of perception systems.

### Definition 10

A perception system  $S$  is called *C-serial* one if for every  $g \in S$ , there is  $g' \in S$  such that  $Acc(g') = Per(g)$  and  $Per(g') = Per(g)$ .

As we assume that the agent is not aware of her faulty sensors, the agent thinks of a global state possible iff the agent's accessible state in the global state is the same as her perception. Therefore, the  $C$ -seriality guarantees that there is always a global state that the agent thinks of possible.

Because the agent has no incorrect observation in a *possible* global states guaranteed by the  $C$ -seriality, we can see

that the  $C$ -seriality implies the directness for any perception system.

Let  $\vdash_{KC+D_C}$  denote the proof system resulting from  $\vdash_{KC}$  by adding  $\vdash_{KC+D_C} \neg C false$ , where  $false$  is proposition of the form  $\neg p \wedge p$  for some primitive proposition  $p$ . We can get that proof system  $\vdash_{KC+D_C}$  is sound and complete with respect to interpreted  $C$ -serial perception systems.

It is interesting to note that in proof system  $\vdash_{KC+D_C}$ , the modality  $C$  can be eliminated by  $\vdash_{KC+D_C} C\varphi \Leftrightarrow \neg K \neg K \varphi$  (recalling Proposition 5). Moreover, the resulting proof system is just S4.2.

On the other hand, if we put  $C false$  into proof system  $\vdash_{KC}$ , then each formula of form  $C\varphi$  can be replaced by a tautology like  $\neg false$ , and hence proof system  $\vdash_{KC}$  degenerates into the standard epistemic logic S4. Similarly, proof systems  $\vdash_{KC+d}$  and  $\vdash_{KC+c}$  may degenerate into the standard epistemic logics S4.2 and S4.3, respectively.

## Related Work

This work is more or less related the issues of awareness (Halpern and Rego 2005), uncertainty (Halpern 2003), and plausibility (Friedman and Halpern 1999). However, as this paper hinges on a grounded semantics for epistemic logics other than S5, we mainly discuss the related work on epistemic logic and computationally grounded semantics.

## Logics of knowledge and certainty

Epistemic logicians have suggested a list of modal logics as epistemic ones. These logics include S4 (Hintikka 1962), S4.4 (Kutschera 1976), S4.2 (Lenzen 1978), and S4.3 (Hoek 1996). They assume some features (axioms) of the logical behavior of epistemic concepts and then construct various kinds possible world semantics that satisfy the pre-assumed feature (axioms). However, the aim of these logics is not to provide a concrete computational approach to evaluating statements in epistemic logic, though philosophical intuitions have been argued subtly for the underlying possible world semantics.

The notion of certainty used in this paper has been first introduced by Lamarre and Shoham (Lamarre and Shoham 1994) and similar notions are Lenzen's *strong belief* (Lenzen 1978) and Voorbraak's *rational belief* (Voorbraak 1991). Lenzen (Lenzen 1978) 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 (Lamarre and Shoham 1994) provide a model theory of knowledge, belief and certainty, with respect to which all Lenzen's collection of axioms are valid; however, their logic is also based on possible world semantics as standard epistemic logics.

## Computationally grounded logics

Besides the interpreted system model, the agent-environment system model (Wooldridge and Lomuscio 2001) is an influential computationally grounded model in the field of agent theory. A multi-modal logic, called  $\mathcal{VSK}$  logic is established as a computationally grounded logic, 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. The perception operator  $\mathcal{S}$  in  $\mathcal{VSK}$  logic corresponds to our certainty operator  $C$ . Intuitively, their interpretation of  $\mathcal{S}\varphi$  is that *the perception received by the agent carries the information  $\varphi$* ; while  $C\varphi$  in this paper means that, according to the received perception, the agent feels certain of  $\varphi$ . Thus, Wooldridge and Lomuscio's notion of perception is an external one, while ours is internal. Another feature of  $\mathcal{VSK}$  logic is that S5 system is adopted for the three modalities.

The deontic interpreted system model (Lomuscio and Sergot 2003) is also an interesting extension of the interpreted system model. The basic idea is to label an agent's accessible state as red ones or green ones. It can be related to the perception system model, 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.

The interpreted  $KBC$  system model is a natural computationally grounded model of agency (Su et al. 2005). A logic of knowledge, belief and certainty, called the KBC logic, is established based on this model. Both the interpreted  $KBC$  system model and the interpreted perception model share the same idea that the sensors of an agent may become inaccurate and thus the visible part of the environment (or the agent's accessible state) may differ from the perception received by the agent. However, the two models are significantly different in that the agent's perception, in the interpreted perception model, is divided into an  $n$ -tuple, which enables us to express which parts of the agent's perception are corrected. Moreover, the statement  $K\varphi$  in the interpreted  $KBC$  system model essentially means that the agent *should know* if she did not have any faulty sensor, while in the interpreted perception system model,  $K\varphi$  indicates that the agent can deduce  $\varphi$  from her correct parts of perception (and her knowledge about the system).

## Conclusions

We have extended the interpreted system model, and provided a computationally grounded model, called the interpreted perception system model, to those epistemic logics other than S5. The key idea is to assume that an agent may have incorrect observations to her accessible state, and we deduce her knowledge by using only her correct parts of the observations. Based on the interpreted perception system model, we have established a logic of knowledge and certainty, which may degenerate into the standard epistemic logic S4 when we eliminate modality  $C$ . We have also define several classes of perception system models, with respect to which, the related proof systems have been shown to be sound and complete, respectively. When only modality  $K$  is considered, the related proof systems become S4.2 and 4.3, two influential epistemic logics in the philosophic community.

We are currently working on those temporal epistemic logics, of which the epistemic dimension is other than S5. It is very natural to introduce temporal operators only if we consider systems of *runs* instead of those of extended global states.

## Acknowledgement

This work was supported by National Natural Science Foundation of China grants 60496327,60725207 and 60473004, 973 Program of China under grant 2005CB321902, and the Australian Research Council grant DP0452628.

## References

- Blackburn, P.; de Rijke, M.; and Venema, Y. 2001. *Modal Logic*. Cambridge University Press.
- Fagin, R.; Halpern, J.; Moses, Y.; and Vardi, M. 1995. *Reasoning about knowledge*. Cambridge, MA: MIT Press.
- Friedman, N., and Halpern, J. Y. 1999. Modeling belief in dynamic systems. part ii: Revision and update. *Journal of AI Research* 10:117–167.
- Halpern, J., and Moses, Y. 1992. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence* 54:319–379.
- Halpern, J. Y., and Rego, L. 2005. Interactive awareness revisited. In *Proc. of TARK-05*, 78–91.
- Halpern, J., and Vardi, M. 1986. The complexity of reasoning about knowledge and time: extended abstract. In *Proc. STOC-86*.
- Halpern, J., and Zuck, L. 1992. A little knowledge goes a long way: Simple knowledge based derivations and correctness proofs for a family of protocols. *Journal of the ACM* 39(3):449–478.
- Halpern, J. Y. 2003. *Reasoning About Uncertainty*. MIT.
- Hintikka, J. 1962. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell: Cornell University Press.
- Hoek, W. v. d. 1996. Systems for knowledge and belief. *Journal of Logic and Computation* 3:173–195.
- Hughes, G., and Cresswell, M. 1996. *A New Introduction to Modal Logic*. London: Routledge.
- Kripke, S. 1963. A semantical analysis of modal logic. i: Normal modal propositional calculi. *Z. Math. Logik Grundl. Math.* 9:67–96.
- Kutschera, F. v. 1976. *Einführung in die intensional Semantik*. Berlin: W. de Gruyter.
- Lamarre, P., and Shoham, Y. 1994. Knowledge, certainty, belief, and conditionalisation. In *KR94*, 415–424. Morgan Kaufmann.
- Lenzen, W. 1978. *Recent Work in Epistemic Logic*, volume XXX of *Acta Philosophica Fennica*. Amsterdam: North-Holland.
- Lomuscio, A., and Sergot, M. 2003. Deontic interpreted systems. *Studia Logica* 75.
- Meyer, J.-J., and van der Hoek, W. 1995. *Epistemic Logic for Computer Science and Artificial Intelligence*. Cambridge Tracts in Theoretical Computer Science 41. Cambridge University Press.
- Su, K.; Sattar, A.; Governatori, G.; and Chen, Q. 2005. A computationally grounded logic of knowledge, belief and certainty. In *Proc. of AAMAS-05*.
- van Benthem, J. 1984. Correspondence theory. In Gabbay, D., and Guenther, F., eds., *Handbook of Philosophical Logic*, volume II. D. Reidel Publ. Co. 167–247.
- Voorbraak, F. 1991. The theory of objective knowledge and rational belief. In *Logics in AI, European Workshop JELIA '90*, 499–515. Springer Verlag.
- Wooldridge, M., and Lomuscio, A. 2001. A computationally grounded logic of visibility, perception, and knowledge. *Logic Journal of the IGPL* 9(2):273–288.
- Wooldridge, M. 2000. Computationally grounded theories of agency. In Durfee, E., ed., *ICMAS-00*, 13–22. IEEE Press.