ICOT Technical Report: TR-0880

TR-0880

On The Semantics of A Shared Common Knowledge Distributed Logic System

by
Wang Xianchang
Yuan Li-Yan (Univ. of Alberta),
You Jia-Huai (Univ. of Alberta)

July, 1994

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03)3456-3191～5

Institute for New Generation Computer Technology

On The Semantics of A Shared Common Knowledge Distributed Logic System

Wang Xianchang

Institute for New Generation Computer Technology (ICOT)

1-4-28, Mita 1-Chome, Minato-ku, Tokyo 108 Japan

Department of Computer Science

Changsha Institute of Technology, Hunan, P.R. China

Yuan Li-Yan, You Jia-Huai

Department of Computing Science

University of Alberta, Edmonton, Canada

First author's present and complete correspondence address:

Dr. Wang Xianchang, STA Fellow

Institute For New Generation Computer Technology (ICOT)

1-4-28, Mita 1-Chome, Minato-ku, Tokyo 108 Japan

Voice: +81-3-3456-3192, Fax: +81-3-3456-1618

e-mail: wang@icot.or.jp

June 28, 1994

Abstract

In this paper, a complete logic system, W, that is based on shared common knowledge views is proposed. The basic properties of W are: 1. Every agent is logically closed. That is, every agent knows the conclusion of its knowledge. 2. One normal agent's knowledge need not necessarily be true in the real world. This means that if i is not a fool agent, it is a normal agent, then $Kip \wedge \neg p$ is consistent. 3. Common knowledge is true in the real world. 4. Common knowledge is the so-called Fool's knowledge, in that every agent knows it, every agent knows that every agent knows it, and so on. 4. Compared with preciously published works on common knowledge, our logic system emphasizes how to use common knowledge, rather than how to define common knowledge. This logical system describes a multi-agent reasoning system based on shared common knowledge views. And, finally, this logical system, W, is sound and complete.

Key Words: Modal, Logic, Knowledge, Common Knowledge, Distributed, Agent.

1 Introduction

In recent years, the representation and reasoning of knowledge and common knowledge have become more and more important topics of research in the field of AI. This is because 'intelligent agents must be able to reason about their own knowledge as well as other agents' knowledge' [20], and 'reasoning about knowledge is also crucial in understanding and reasoning about protocols in distributed systems' [7].

The building of such a logic system, and fitting it to distributed processing systems is not an easy task. Many knowledge-based multi-agent logic systems have been proposed, but all suffer from their own particular problems, because 'there is no agreement on exactly what the properties of knowledge are or should be. For example, is it the case that you know what facts you know? do you know what you don't know? do you know only true things, or can something you 'know' actually be false?' [7].

Generally, the main disagreements related to the properties of knowledge-based multi-agent reasoning systems are as follows:

1. Should each normal agent's knowledge be true in the real world? That is, should

the statement $Kip \to p$ (i is not a fool) be true in the knowledge system? According to some authors' arguments, if a logic system does not contain the axiom $Kip \to p$, then the logical system is called belief system.

2. Should real world knowledge be known by all agents? That is: should the necessity inference rule of the classic modal logic, $p \Rightarrow Kip$ be included in the knowledge system? Many interesting multi-agent reasoning puzzles, for example, the Conway Paradox puzzle [22], and the Three Wise Men puzzle [10] [23], require that real world knowledge not be known by any agent.

3. Should the knowledge distributed axioms $Ki(p \to q) \to (Kip \to Kiq)$ be held in the logic system? Generally, if we do not consider the logical omniscience problem [8], then we generally accept this axiom. In fact, most knowledge and belief logic systems accept this axiom. Such logical systems are referred to as normal modal logic in [16].

4. Should an agent have positive introspective ability? That is, should the axiom $Kip \to KiKip$ be included in the logic system? Most researchers agree that this axiom should be included.

5. Should the agent have negative introspective ability? That is, should the axiom $\neg Kip \to Ki\neg Kip$ be included in the logic system? There is great disagreement on this point. This axiom has a little relevance to the necessary rule. If we take $\neg Kip$ as real world knowledge, which means that an outsider (or a god) can observe that agent i does not know statement p, then, according to this axiom, agent i should know that he does not know statement p. In our opinion, this is too strong a property to assign to an agent, therefore in our logic, we do not accept this axiom.

6. If we suppose an agent to be consistent, then it is reasonable to believe that if agent i knows $\neg p$, then agent i will not know p. So, we believe that an idea logic system will contain the axiom $Ki\neg p \to \neg Kip$.

Briefly, our opinions about a multi agent logic system are as follows:

1. Tautology should be known by any agent, every tautology being decidable by every agent.

   For example, suppose that i is an agent, and p is a statement. Then, $Ki(p \vee \neg p)$ is true and agent i can prove that $p \vee \neg p$ is true.

2. The agent's knowledge need not be complete.

   That is, some knowledge p and its negative $\neg p$ will not be known by agent i. So $Kip \vee Ki\neg p$ should not be a conclusion of our logical system.

3. The knowledge known by a normal agent i, can be inconsistent with the real world.

   This means that, suppose i is a normal agent, $(Kip) \wedge \neg p$ is consistent in our logic system [ref Example 3.1]. Then, the axiom in modal logic S5, $Kip \rightarrow p$ can not be held in our logic system.

4. Real world knowledge should not be known by any agent. This means that the necessitation rule in modal logic can not be included in our logical system.

   One of the basic ideas behind our logic system is: a true real world knowledge 'p' is not necessarily obvious to every agent. In other words, real world knowledge is not shared common knowledge.

5. Every agent should be positive introspective, and should not be negative introspective.

6. Considering the real world knowledge and agent's knowledge, we can assume that if agent knows $\neg p$ then that agent should not know p. That is, if an outsider (or a god) observes that agent i knows $\neg p$, then the outsider will assume that agent i does not know p.

7. Common knowledge should be typical knowledge of our system. Common knowledge should be black board knowledge. In most preciously published papers such as: [1][7][6], it was assumed that common knowledge should be defined by infinite deductions. Our opinion about common knowledge is that, common knowledge should have the infinite deductive properties, but it should not be limitted to and defined by this property.

   Common knowledge should have the following properties:

   (a) First, as in [15], we introduce a fool reasoner and assume that whatever a fool knows is common knowledge. In this paper we assume that 0 is the fool. Then, if K0p appears in a theory, p should be common knowledge.

   (b) Tautology should be common knowledge.

   (c) Common knowledge should be true in the real world. This means that we should accept the axiom $K0p \rightarrow p$ and the safeness rule $K0p \Rightarrow p$.

   Further, if we consider the fact that common knowledge should be true in the real world, is also a common knowledge, that means , if we accept $K0(K0p \rightarrow p)$ as an axiom, then we should also have the safe rule $K0p \Rightarrow p$ be included in our logical system.

(d) If p is common knowledge, then for every agent i, Kip should also be common knowledge.

(e) The Fool should have positive introspective ability. This means that, if p is common knowledge, then $K0p$ will also be common knowledge.

(f) Compared with the common knowledge definition in [1], one of the most important aspects of our logic system should be: We take care only of how to use common knowledge rather than concerning ourselves with what common knowledge is.

In the past few years, we have tried to devise a logic system that describe the above properties. This paper presents an improved and complete version that is based on our findings [22][23][24][25]. Main reasons undertaking this project are:

1. Previously, there were no logic systems that could satisfactorily reflect the typical properties of a shared common knowledge distributed knowledge system.

2. Classic and improved modal logic can not (at least, not without great difficulty) deal with distributed knowledge processing systems, there was, therefore a pressing need to devise a satisfactory modal logic.

3. Because distributing processing systems are becoming more and more important as the scale of knowledge processing increases, it can be clearly seen that there is an urgent need to find a satisfactory logic model for distributed processing.

Figure 1 simply illustrates a multi-agent system that is based on shared common knowledge views in which $\tilde{p}$ is same as $\neg p$. It includes the real world, the fool agent, the normal agents, and shows how each is interrelated.

The paper is organized as follows: In section 2, we present a formal logic system, W, which is an improvement of our logic system presented in [22][23]. W has all the meta-properties of classic modal logic, including compactness, as well as monotonic and deductive properties. Important results are: Every agent's knowledge is closed under W [ref Theorem 2.7]; If a theory T's *-translation is consistent, then T is also consistent [ref Theorem 2.10]. In section 3, we present the model theory on which W is explained. We propose W-Kripke structures, based on the Kripke structure, and provide some explanations about the W-Kripke structure by means of examples. In section 4, we prove W's soundness. In section 5, we prove W's completeness, based on the Canonical W-Kripke structure. Finally, in our conclusion, we briefly review our work and future work to be performed regarding W.

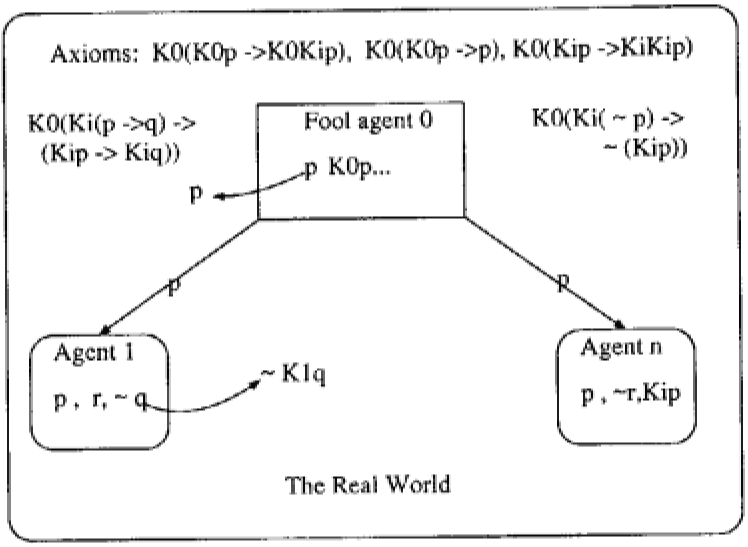


Figure 1: A Model for Shared Common Knowledge Multi-agent Reasoning System

2 W System

Suppose At is a set of primitive statements. $Ag = \{0, 1, ..., n\}$ is the set of agent, in which 0 is called the fool agent, the rest is called the normal agent or agent if it is not confused. Informally, 0's knowledge is common knowledge, which is known to all agents.

First, we define the syntax of the well-founded formulas based on At and Ag.

Definition 2.1 A well-founded formula based on At and Ag can be inductively defined as follows:

1. If $p \in At$, then p is a well-founded formula.

2. If p, q are well-founded formulas, $i \in Ag$, then Kip, $(\neg p), (p \rightarrow q)$ are also well-founded formulas.

3. All well-founded formulas are defined by the finite compositions of steps 1 and 2.

□

We denote the set of all the well-founded formulas based on At and Ag, by L.

We use special symbols to abbreviate some formulas. We write $(p \vee q)$ for $(\neg p \rightarrow q)$, $p \wedge q$ for $\neg(p \rightarrow \neg q)$, $p \equiv p$ for $(p \rightarrow q) \wedge (q \rightarrow p)$. Assume formula P to be a basic formula if P contains no modal operator.

The axioms and inference rules of W are defined as shown below. This is an improved logic system, based on [22] [23].

Definition 2.2 W's axioms:

A1. $K0p$, if p is any tautology.

A2. $K0(K0p \rightarrow K0Kip)$.

A3. $K0(Ki(p \rightarrow q) \rightarrow (Kip \rightarrow Kiq))$.

A4. $K0(K0p \rightarrow p)$.

A5. $K0(Kip \rightarrow KiKip)$.

A6. $K0(Ki\neg p \rightarrow \neg Kip)$.

W's inference rules are

Modus Ponens: $p, p \rightarrow q \Longrightarrow q$

Safeness rule: $K0p \Longrightarrow p$ □

Notice that, safeness rule can be excluded from W, if we add a new axiom $K0p \rightarrow p$ into W.

Definition 2.3 Extension

Suppose theory $T \subseteq L$, we define $Cons(T)$, the extension of T, as being the smallest subset of L that satisfies the following conditions:

1. $T\cup$ Axioms $\subseteq Cons(T)$
2. If $p \in Cons(T), p \rightarrow q \in Cons(T)$ then $q \in Cons(T)$
3. If $K0p \in Cons(T)$ then $p \in Cons(T)$ □

Obviously, the concept of extension is well-defined and unique for every theory.

Now, it is not difficult to prove the following theorem.

Theorem 2.1 *Constructive Property of Cons(T).*

Suppose T is a theory. We can inductively construct the following sets:

$Cons_0(T) = Axioms \cup T$, *and for all* $i \geq 0$:

$Cons_{i+1}(T) = Cons_i(T) \cup \{q|$ *there are formula p such that* $\{p, p \rightarrow q\} \subseteq Cons_i(T)$, *or* $K0q \in Cons_i(T)\}$

Then $Cons(T) = \sum_{i=0}^{i=\infty} Cons_i(T)$ □

Suppose T is a theory. As in [2], we can define the prove relationship between T and well-formed formula p. We denote this by $T \vdash p$, where p is called the consequence of T. Obviously, the consequence set of T is $Cons(T)$. that is, $Cons(T) = \{p|T \vdash p\}$.

Definition 2.4 [Contradictory]

Say theory T contradicts the agent sequence $< i_1, ..., i_k >$, if there is a formula p, such that both p and $K_{i_1}...Ki_k \neg p$ can be proven under T. If $k = 0$, say T is contradictory (or inconsistent). If $k = 1$, say T is contradictory about agent i_1. □

Say theory T is consistent, if it is not contradictory. Say T is consistent about agent sequence $< i_1, ..., i_k >$, if it is not contradictory about $< i_1, ..., i_k >$. Obviously, if theory T is consistent about agent sequence $< i_1, ..., i_k >$, then it must be consistent.

If theory T is not consistent, then for every well-formed formula p, we have $T \vdash \neg(p \rightarrow p)$.

Theorem 2.2 *Compactness theorem*

$p \in Cons(T)$ iff there is a finite subset T' of T, such that $p \in Cons(T)$ Or equally $Cons(T) = \bigcup\{Cons(T') | T' \subseteq T$ and T' is finite $\}$ □

Suppose T, T' are two sets of formulas. We then write $T \vdash T'$ as the abbreviation: for every $p \in T', T \vdash p$. From the compactness theorem, we can easily get:

Corollary 2.3 Monotonicity of W.

Suppose $T1, T2, T3$ are sets of formulas, if $T1 \vdash T2, T2 \vdash T3$ then $T1 \vdash T3$. □

Theorem 2.4 *Deduction theorem*

Suppose T is a theory, p, q are two formulas, then $T \bigcup\{p\} \vdash q$ if and only if $T \vdash p \rightarrow q$. □

Lemma 2.1 Suppose p is a tautology, i1,...,in are agents, then $\vdash K_{i1}...K_{in}p$ holds. □

Whether this lemma holds depends only on axioms A1, A2. It shows that every agent knows the tautology, and that every agent knows that other agents (include itself) know the tautology, and so on.

Lemma 2.2 For every formula p,q, agent i, $Ki(p \rightarrow q) \vdash (Kip \rightarrow Kiq)$. □

Whether this lemma holds depends only on axiom A3. Notice that this lemma does not mean $(Kip \rightarrow Kiq) \vdash Ki(p \rightarrow q)$. Generally we do not have $(Kip \rightarrow Kiq) \vdash Ki(p \rightarrow q)$.

Lemma 2.3 For every formula p, $p \in Cons(\{\})$ iff $K0p \in Cons(\{\})$. □

Whether this lemma holds depends only on the following two assumptions:

1. Every axiom of W is in the form of $K0(...)$;
2. Knowledge distributed axiom A3 is needed.

Notice that this lemma does not mean $p \vdash K0p$.

Corollary 2.5 For every agent i, if $p \in Cons(\{\})$ then $Kip \in Cons(\{\})$. □

Theorem 2.6 *For every well-formed formula p, q, agent i, we have*
$\vdash K0(Kip \wedge K0(p \rightarrow q) \rightarrow Kiq)$ *and*
$\vdash K0(K0p \wedge Ki(p \rightarrow q) \rightarrow Kiq)$ □

This theorem shows that the fact that every agent can do modus ponens reasoning based on its own knowledge and common knowledge is, itself, common knowledge.

Theorem 2.7 *Suppose T is a theory and that $T = Cons(T)$. For any agent $i \in Ag$, let $T/Ki = \{p|Kip \in T\}$, then $T/Ki = Cons(T/Ki)$.* □

Proof:

Suppose $p \in Cons(T/Ki)$, then, according the compactness and deductive theory, there are some formulas $\{p1, ..., pn\} \subseteq T/Ki$, such that $\vdash (p1 \rightarrow (... \rightarrow (pn \rightarrow p)...))$. According to Corollary 2.5 and Lemma 2.2, we get $\vdash (Kip1 \rightarrow (... \rightarrow (Kipn \rightarrow Kip)...))$. Since $\{Kip1, ..., Kipn\} \subseteq T$, $T = Cons(T)$, so $Kip \in T$. So $p \in T/Ki$.

This theorem shows that every agent's knowledge is logical closed. That is, every agent has the same inference ability as W.

Corollary 2.8 Suppose p1, ..., pn, q are well-formed formulas, and that $i_1, ..., i_k$ are agents. If $p1, ..., pn \vdash q$, then $K_{i_1}...K_{i_k}p1, ..., K_{i_1}...K_{i_k}pn \vdash K_{i_1}...K_{i_k}q$. □

Proof:

Suppose $T = Cons(\{K_{i_1}...K_{i_k}p1, ..., K_{i_1}...K_{i_k}pn\})$, $T' = T/K_{i_1}.../K_{i_k}$. Since $\{p1, ..., pn\} \subseteq T'$, $p1, ..., pn \vdash q$. So, according to Theorem 2.7, $q \in T'$. Hence, $K_{i_1}...K_{i_k}q \in T$. So $K_{i_1}...K_{i_k}p1, ..., K_{i_1}...K_{i_k}pn \vdash K_{i_1}...K_{i_k}q$.

From this corollary, if q is a propositional logical consequence of formulas p1, ..., pn, then the above statement does also hold.

Corollary 2.9 Suppose T is a theory, p is a formula, if $T \nvdash p$ then $T \bigcup \{\neg p\}$ is consistent.

Proof:

Suppose $T \bigcup \{\neg p\}$ is not consistent, then $T \bigcup \{\neg p\} \vdash \neg(p \rightarrow p)$. By applying the compactness theory, we get $T \vdash \neg p \rightarrow (\neg(p \rightarrow p))$, so $T \vdash (p \rightarrow p) \rightarrow p$, so we get $T \vdash p$, which contradicts the assumption $T \nvdash p$. So, $T \bigcup \{\neg p\}$ must be consistent. □

Definition 2.5 Suppose P is a modal formula. We define P^*, P's $*$-translation, as a formula that contains no modal operator. P^* is defined inductively as follows:

1. If P is a basic formula, then P^*=P.
2. $(P \rightarrow Q)^* = (P^* \rightarrow Q^*)$ and $(\neg P)^* = \neg P^*$.
3. $(KiP)^* = P^*$ □

For example, suppose p, q are two basic formulas, then $((Ki(p \rightarrow q) \rightarrow (Kip \rightarrow Kiq))^* = (p \rightarrow q) \rightarrow (p \rightarrow q))$.

Lemma 2.4 Suppose formula $P \in Cons(\{\})$, then P^* is a tautology. □

Whether this lemma holds depends only on the fact that every axioms' $*$-translation is a tautology.

Theorem 2.10 *Suppose T is a finite theory, if $T*$ is consistent under propositional calculus, then T is consistent about every agent sequence $< i1, ..., ik >$.* □

Proof:

Suppose T^* is consistent under proposition calculus. We can prove that T is consistent about every agent sequence $< i1, ..., ik >$ $(k \geq 0)$. If T is not $< i1, ..., ik >$ consistent. Then, there must be a statement p, such that $p \wedge Ki1...Kik(\neg p) \in Cons(T)$. According to the deduction theorem and Lemma 1.4, we have $T \rightarrow (p \wedge Ki1...Kik(\neg p)) \in Cons(\{\})$, and $T* \rightarrow (p * \wedge \neg p*)$ is a tautology. So $\neg T*$ must be a tautology, which contradicts the assumption that $T*$ is consistent under propositional calculus. So, T is consistent about every agent sequence $< i1, ..., ik >$.

Corollary 2.11 $Cons(\{\})$ is consistent about every agent sequence $< i1, ..., ik >$. □

Generally, a theory T can be divided into $n+2$ parts. $T = T_r \cup T_0 \cup T_1 \cup ... \cup T_n$, Where for $i \in Ag$, T_i denote all the $Ki(...)$ formulas in T, T_r is the rest formulas of T. Some times, we denote T_i by $Ki\{p | Kip \in T_i\}$. We call p is common knowledge, if $K0p \in T_0$; p is agent i's knowledge if $Kip \in T_i$; else p is regarded as real world knowledge.

Example 2.1 [Conway Paradox]

During a card game, both Max and Pat have an ace. If either is asked whether they have any knowledge about the other person's cards they will answer 'no'. Their answer will not change if the question is repeated. But, if someone tells them "at least one of you has an ace", a fact they can infer from their own cards, the answer will be 'no' the first time they are answered(Max), and 'Yes, he/she has an ace' the second time (Pat). How can we deal with this kind of reasoning? □

Suppose p is the statement 'Max has an ace', and q is ' Pat has an ace'. Then, the Conway Paradox 's formal description as a theory is follows:

$T0 = \{p, q, K0(p \rightarrow K1p), K0(\neg p \rightarrow K1\neg p),$

$K0(q \rightarrow K2q), K0(\neg q \rightarrow K2\neg q), K1p, K2q\} =$

$\{p, q\}\vee$;system knowledge

$K0\{p \rightarrow K1p, \neg p \rightarrow K1\neg p, q \rightarrow K2q, \neg q \rightarrow K2\neg q\}\vee$;common knowledge

$K1\{p\}\vee$;agent 1 (Max)'s knowledge

$K2\{q\}$;agent 2 (Pat)'s knowledge

According to theory T0, we can not obtain any new information by repeating the inquiring. That is, T0 is consistent with the formulas set Rp=$\{K0\neg K1q, K0\neg K2p\}$.

Now, if we tell the players "at least one of you has an ace", which means if we add $K0(p \vee q)$ into T0, and get theory $T1 = T0 \cup \{K0(p \vee q)\}$. Then, after we add the first answer $K0\neg K1q$ to T1, and get theory T2, then we can conclude $K2p$ from T2 [1]. The proof is as follows:

$T2 \vdash K2p$

01. $K0(K1\neg p \wedge K0(\neg p \rightarrow q) \rightarrow K1q)$ ......Theorem 2.6
02. $K0\neg K1q$ T2
03. $K0(\neg K1\neg p \vee \neg K0(\neg p \rightarrow q))$ ......1, 2, corollary 2.8
04. $K0(p \vee q)$ T2
05. $K0(\neg p \rightarrow q)$ ......4, corollary 2.8
06. $K0K0(\neg p \rightarrow q)$ ......5, axiom 1, and safeness rule
07. $K0\neg K1\neg p$ ......03, 06 and corollary 2.8
08. $K0(\neg p \rightarrow K1\neg p)$ T2
09. $K0p$ ......7,8, and corollary 2.8
10. $K0K2p$ ......9, axiom 1

[1] In fact, to prove $K2p$, we should first prove $K0p$, meaning that, under theory T2, even the fool can conclude that p is true.

11. $K2p$ ......10, safeness rule

Hence $T2 \vdash K2p$.

From Theorem 2.10, we can prove that T1 is consistent because $T1^*$ is consistent.

To prove T2's consistency, we have to find a model for T2. This is given in the next section.

3 Model Theory: W-Kripke Structure

Definition 3.1 W-Kripke Structure

Suppose L is a language based on At and Ag. $\kappa = (W, \pi, w0, R0, R1, ..., Rn)$ is a Kripke structure based on L, where W is an non-empty set, called the world set. $w0 \in W$ is called an initial world; π is a map from W to the subset of At; R0,R1,...,Rn are relations on W. Say structure κ is a W-Kripke structure, if κ satisfies the following four conditions:

1. Every Ri ($i = 0, 1, ..., n$) is transitive;
2. For every $i = 1, ..., n$, $Ri \subseteq R0$;
3. R0 is reflexive;
4. Every Ri is serial. That is, for every world $w \in W$, every agent $i \in Ag$, the set $\{w' | (w, w') \in Ri\}$ is not empty. □

Generally, we denote id for the reflexive relation on W, $id = \{(w, w) | w \in W\}$.

Definition 3.2 Suppose $\kappa = (W, \pi, w0, R0, R1, ..., Rn)$ is a W-Kripke structure. We define the semantics entailment relation $\kappa, w \models q$, as follows:

1. If $p \in At$, then $\kappa, w \models p$ iff $p \in \pi(w)$
2. $\kappa, w \models \neg p$ iff $\kappa, w \not\models p$
3. $\kappa, w \models p \rightarrow q$ iff if $\kappa, w \not\models p$ or $\kappa, w \models q$
4. For every $i \in Ag$, $\kappa, w \models Kip$ iff for every $w' \in W$, if $(w, w') \in Ri$, then $\kappa, w' \models p$

□

Definition 3.3 Suppose T is a theory, p is a formula, for every W-Kripke structure $\kappa = (W, \pi, w0, R0, R1, ..., Rn)$, we define:

Formula p is valid in κ, denoted as $\kappa \models p$, if $\kappa, w0 \models p$;

Say theory T is valid in W-Kripke structure κ, denoted as $\kappa \models T$ if, for every formula $p \in T$, p is valid in κ;

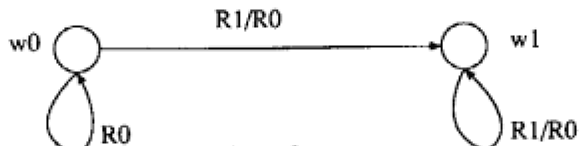


Figure 2

Say formula p is a semantic entailment of theory T, denoted as $T \models p$, if for every T's valid W-Kripke structure κ, p is also valid in κ;

We denote the set of all the semantic entailment of theory T by Th(T). □

Example 3.1 Suppose $At = \{p\}$, $Ag = \{0, 1\}$, then theory $T = \{p, K1\neg p\}$ can be satisfied. One of T's W-Kripke model is $\kappa = (W, \pi, w0, R0, R1)$:

$W = \{wo, w1\}$, $\pi(w0) = \{p\}, \pi(w1) = \{\}$, $R0 = \{(w0, w0), (w0, w1), (w1, w1),$ $R1 = \{(w0, w1), (w1, w1)\}$.

Above W-Kripke model can be described by Figure 2, where every node is the world, the link denoted by Ri from node w to w' express that $(w, w') \in Ri$. □

Example 3 is adapted from [20] [5].

Example 3.2 Suppose $At = \{p\}$, $Ag = \{0, 1, 2\}$ where 1 is agent Alice and , 2 is agent Bob. The statements in [20] [5] are:

p is true;

Alice doesn't know whether p is true or false;

Bob knows that p is true;

Alice knows that Alice doesn't know about p, but Alice knows that Bob knows whether p is true or false;

Bob knows that he knows p, but he doesn't know whether Alice knows p;

Alice knows that Bob doesn't know whether Alice knows about p.

The above statements can be described by theory T as:

$T = \{p, \neg K1p, \neg K1\neg p, K2p, K1\neg K1p, K1\neg K1\neg p,$
$K1(K2p \vee K2\neg p), K2K2p, \neg K2(K1p \vee K1\neg p),$
$K1\neg K2(K1p \vee K1\neg p)\}$

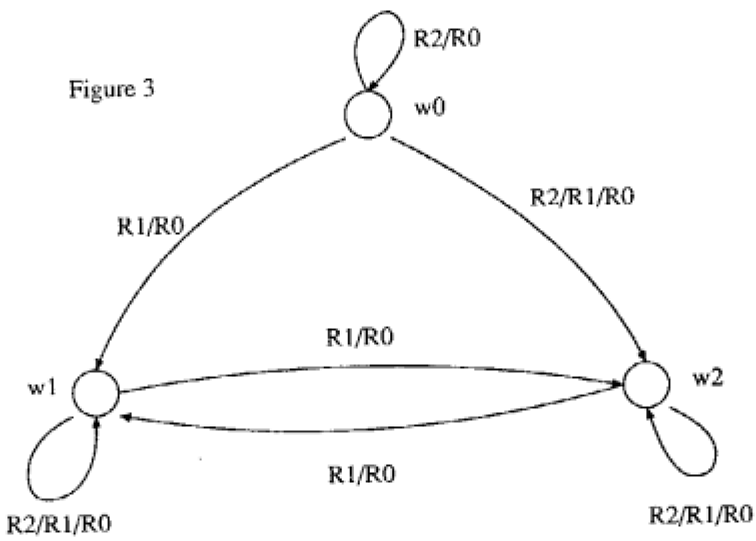
one of T's W-Kripke models $\kappa = (W, \pi, w0, R0, R1, R2)$ is:

$W = \{w0 = \{p\}, w1 = \{\}, w2 = \{p\}\}$

$R1 = \{(w0, w1), (w0, w2), (w1, w1), (w1, w2), (w2, w2)\}$

$R2 = \{(w0, w0), (w0, w2), (w2, w2), (w1, w1)\}$

Figure 3



$R0 = id \cup R1 \cup R2$.

κ can be described as shown in Figure 3. □

Example 3.3 Let's consider Example 2.1 in the previous section.

$T0 = \{p, q, K0(p \rightarrow K1p), K0(\neg p \rightarrow K1\neg p),$

$K0(q \rightarrow K2q), K0(\neg q \rightarrow K2\neg q), K1p, K2q\}$ has the following W-Kripke structure models:

$\kappa 0 = (W0, \pi, w0, R0, R1, R2)$

$W0 = \{w0 = \{p, q\}, w1 = \{p\}, w2 = \{q\}, w3 = \{\}\}$

$R1 = \{(w0, w1), (w1, w1), (w2, w2), (w2, w3), (w3, w3)\}$

$R2 = \{(w0, w2), (w2, w2), (w1, w1), (w1, w3), (w3, w3)\}$

$R0 = id \cup R1 \cup R2 \cup \{(w0, w3)\}$

We can see that $\neg K0(p \vee q)$, $K0\neg K1q$, $K0\neg K2p$, $K1\neg K0(p \vee q)$, $K2\neg K0(p \vee q)$, K2q, $K2\neg K1p$, $K1\neg K2q$ are satisfied in model $\kappa 0$.

$\kappa 0$ can be described by Figure 4.

$T1 = T0 \cup \{K0(p \vee q)\}$ has the following W-Kripke structure model:

$\kappa 1 = (W1, \pi, w0, R0, R1, R2)$

$W1 = \{w0 = \{p, q\}, w1 = \{p\}, w2 = \{q\}\}$

$R1 = \{(w0, w0), (w0, w1), (w1, w1), (w2, w2)\}$

$R2 = \{(w0, w0), (w0, w2), (w2, w2), (w1, w1)\}$

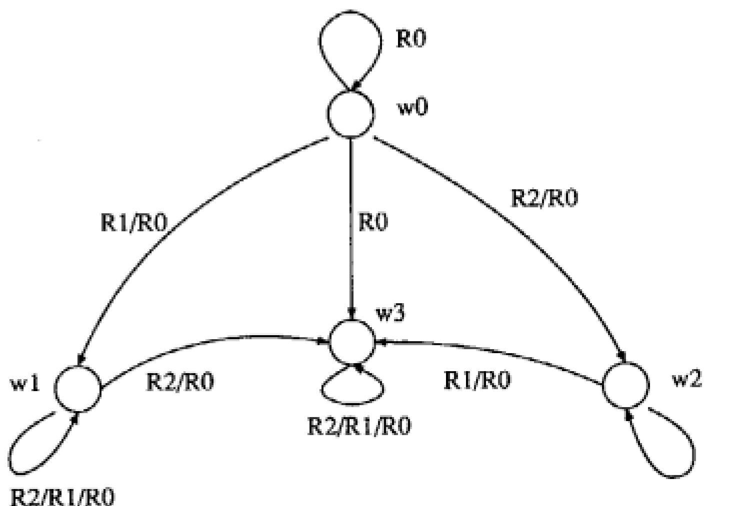


Figure 4

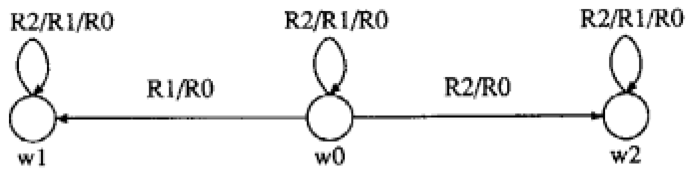


Figure 5

$R0 = id \cup R1 \cup R2$

We can see that $K0(p \vee q)$, $\neg K0 \neg K1q$, $\neg K0 \neg K1 \neg q$, $\neg K0 \neg K2p$, $\neg K0 \neg K2 \neg p$, $K0 \neg K1q \rightarrow K0q$ are satisfied in model $\kappa 1$.

$\kappa 1$ can be described by Figure 5.

$T2 = T1 \cup \{K0 \neg K1q\}$ has the following W-Kripke structure model:

$\kappa 2 = (W2, \pi, w0, R0, R1, R2)$

$W2 = \{w0 = \{p, q\}, w1 = \{p\}\}$

$R1 = \{(w0, w0), (w0, w1), (w1, w1)\}$

$R2 = \{(w0, w0), (w1, w1)\}$

$R0 = id \cup R1 \cup R2$

In this model, formulas $K0(p \vee q)$, $K0 \neg K1q$, $K1p$, $\neg K1q$, $\neg K1 \neg q$, $K2q$, $K2p$ are satisfied.

Figure 6

$\kappa 2$ can be described by Figure 6.

□

4 W's Soundness

In this section, we consider the soundness of logic system W.

Proposition 4.1 Every axiom in W is valid.

Proof:

Suppose κ is a W-Kripke structure, and w0 is the initial world.

Obviously, A1 is valid.

Now, we can prove that A2 is valid. That is $\kappa, w0 \models K0(K0p \rightarrow K0Kip)$. For every w', if $(w0, w') \in R0$, we should prove κ, $w' \models (K0p \rightarrow K0Kip)$. Suppose κ, $w' \models K0p$. We should prove κ, $w' \models K0Kip$.

First we have, for every w'', if $(w', w'') \in R0$, then κ, $w'' \models p$. Now we need to prove κ, $w'' \models Kip$. That is, for every w''', if $(w'', w''') \in Ri$, then κ, $w''' \models p$. Because $Ri \subset R0$, R0 is transitive, so κ, $w''' \models p$, so κ, $w' \models K0Kip$. So, $\kappa, w' \models (K0p \rightarrow K0Kip)$. Hence $\kappa, w0 \models K0(K0p \rightarrow K0Kip)$.

A3 is valid because $\kappa, w \models p \rightarrow q$ iff $\kappa, w \not\models p$ or $\kappa, w \models q$.

A4 is valid because R0 is reflexive.

A5 is valid because every relation $Ri, i = 0, 1, ..., n$ is transtive.

A6 is valid because every Ri is serial. □

Proposition 4.2 W's inference rules are also safe.

Proof:

Suppose p, $p \rightarrow q$ are true in the W-Kripke structure κ, $\kappa, w0 \models p$ and $\kappa, w0 \models p \rightarrow q$. Then, it is easy to prove that $\kappa, w0 \models q$.

Suppose $\kappa, w0 \models K0p$. Since $(w0, w0) \in R0$, so $\kappa, w0 \models p$. □

From the above two propositions, we can consequently derive W's soundness.

Lemma 4.1 Suppose T is a theory, and κ is T's valid W-Kripke structure. If $p \in Cons(T)$, then $\kappa, w0 \models p$. □

Theorem 4.3 *Soundness of W.*

1. $Cons(\{\})$ is valid. That is, $Cons(\{\}) \subseteq Th(\{\})$.

2. For every theory T, we have $Cons(T) \subseteq Th(T)$ □

Can we have completeness of W? That is, $Cons(\{\}) = Th(\{\})$ and for every theory T, $Cons(T) = Th(T)$. From the above discussion, we already have $Cons(T) \subseteq Th(T)$. In the next section, therefore, we will prove $Cons(T) \supseteq Th(T)$.

5 W's Completeness

In this section, we will prove that the W logic system is complete. The proof demands the application of some special techniques. First, we give the following concepts.

Definition 5.1 Say theory T is complete, if $T = Cons(T)$ and for every formula $p \in L$, either $p \in T$ or $\neg p \in T$. Obviously, L is a complete theory. □

Lemma 5.1 Suppose theory T is consistent, if $p \notin Cons(T)$ then $\{\neg p\} \cup T$ is also consistent.

Proof:

If $\{\neg p\} \cup T$ is not consistent, then $\{\neg p\} \cup T \vdash \neg(p \rightarrow p)$. So $T \vdash \neg p \rightarrow \neg(p \rightarrow p)$. Since $\neg p \rightarrow \neg(p \rightarrow p) \vdash (p \rightarrow p) \rightarrow p$ and $\vdash (p \rightarrow p)$, so we have $T \vdash p$. This contradicts the assumption $p \notin Cons(T)$, so $\{\neg p\} \cup T$ is consistent. □

Theorem 5.1 *Suppose T is a theory, then*

1. If T is inconsistent, then $Cons(T)$ is complete.

2. If T is consistent, then T must have a consistent complete superset theory.

3. If T is consistent complete, then for every agent i, there must be a consistent complete set T' such that $T/Ki \subseteq T'$. □

Proof:

1. Since T is inconsistent, such that there is a formula p such that $\neg p \in T, p \in T$. Since every formula q is a logical consequence of $p \wedge \neg p$, therefore every formula $q \in Cons(T)$. So, T is complete.

2. Suppose T is consistent. Now we prove that T has a consistent complete superset T'. T' is constructed as follows:

   Suppose p1, p2,... is the enumeration of all the formulas of L.

   T_0=T, for every $i \geq 0$, we define

   $T_{i+1} = Cons(T_i)$ if $pi \in Cons(T_i)$ or $\neg pi \in Cons(T_i)$

   $T_{i+1} = Cons(T_i) \cup \{\neg pi\}$ else

   Suppose $T' = \bigcup_{i=0}^{i=\infty} T_i$, then it is easy to prove that T' is a consistent complete superset of T.

3. Suppose $T1 = T/Ki$. It is easy to prove that T1 is also consistent. If T1 is not consistent, then there must be some formulas $\{p1, ..., pn\} \in T1$ such that $\{p1, ..., pn\} \vdash \neg(p \rightarrow p)$. So, we can prove that $\vdash (Kip1 \rightarrow ... \rightarrow (Kipn \rightarrow Ki(\neg(p \rightarrow p))...)$. Since $Ki(\neg(p \rightarrow p)) \vdash \neg Ki(p \rightarrow p)$, we have $\vdash \neg Kip1 \vee ... \vee \neg Kipn$. Notice that $Kip1 \in T, ..., Kipn \in T$, T is not consistent. This is a contradiction. So, T1 is consistent. According to item 2 of this corollary, we can conclude that T1 has a consistent complete superset T', such that $T/Ki \subseteq T'$.

Corollary 5.2 Suppose that T is a consistent theory. Then, for any formula p, if $p \notin T$ and $\neg p \notin T$, T must have a consistent complete superset T' such that $\neg p \in T'$.

Hint: By supposing that the formula p1 in the enumerate sequence of the above theorem is q, then we can obtain this corollary. □

Now, we can construct the Canonical W-Kripke structure, based on consistency theory T, as follows:

Definition 5.2 Canonical W-Kripke structure [ref [16] [6]]

Suppose that T is a consistent theory. We construct the Canonical W-Kripke structure $\kappa = (W, \sigma, w0, R0, ..., Rn)$ as follows:

1. $W = \{T' | T'$ is a consistent complete set$T\}$
2. $w0 \in W$ is a consistent complete superset of T.
3. For every $w \in W$, we define $\sigma(w) = \{p | p \in At$ and $p \in w\}$
4. For every $w1 \in W, w2 \in W$, $(w1, w2) \in Ri$ iff $w1/Ki \subseteq w2$. □

Then it is easy to prove:

Proposition 5.3 κ is a W-Kripke structure.

Proof:

1. Every Ri is transitive.

   If $\{(w1, w2), (w2, w3)\} \subseteq Ri$, then we should be able to prove $(w1, w3) \in Ri$. Suppose $Kip \in w1$, since $w1 = Cons(w1)$, so $KiKip \in w1$. Since $(w1, w2) \in Ri$, so $Kip \in w2$. Since $(w2, w3) \in Ri$, so $p \in w3$. So $(w1, w3) \in Ri$.

2. For agent $i \in Ag$, $Ri \subseteq R0$.

   Obviously, if $(w1, w2) \in Ri$ then, for every formula p, if $Kip \in w1$, then $p \in w2$. Now, we can prove that if $K0p \in w1$ then $p \in w2$. Since $K0p \in w1$ and $w1 = Cons(w1)$, so $Kip \in w1$, so $p \in w2$. So Ri is a subset of R0.

3. R0 is reflexive.

   This is obvious.

4. For every agent i and every world w, the set $\{w'|(w, w') \in Ri\}$ is not empty.

   This is true according to Theorem 5.1(3).

□

Proposition 5.4 For every formula $p \in L$, $\kappa, w \models p$ iff $p \in w$. □

It's proof requires the application of the following lemma:

Lemma 5.2 Suppose w is a consistent complete theory, and p is a formula. If $Kip \notin w$, then there must be a consistent complete theory w', such that $(w, w') \in Ri$ and $\neg p \in w'$.

Proof:

According to Theorem 2.7, suppose $T1 = \{q|Kiq \in w\}$. Since $Kip \notin w$, w is closed, so $p \notin T1$. So, T1 is closed and consistent. According to Corollary 5.2, T1 has a consistent complete super set w', such that $\neg p \in w'$ and $(w, w') \in Ri$. □

Now, it is easy to prove our main statement:

For every formula $p \in L$, $\kappa, w \models p$ **iff** $p \in w$.

This proof is based on the induction of formula p's length.

1. If $p \in At$, then it is obvious that $\kappa, w \models p$ iff $p \in w$.

2. Suppose the above statement is true for every formula p whose length is not greater than t.

3. Suppose p is a formula whose length is greater than t. Then, we can prove the above statement by applying the following.

   (a) p is $\neg q$, where q's length is not greater than t. Then,
   $\kappa, w \models p$ iff $\kappa, w \not\models q$ iff $q \notin w$ iff $\neg q \in w$ iff $p \in w$.

   (b) p is $q \rightarrow r$, where both formula q and r's length are not greater than t.
   $\kappa, w \models p$ iff $\kappa, w \models q \rightarrow r$ iff $\kappa, w \not\models q$ or $\kappa, w \models r$ iff $q \notin w$ or $r \in w$ iff $\neg q \in w$ or $r \in w$ iff $q \rightarrow r \in w$ iff $p \in w$.

   (c) p is Kiq, where q's length is not greater than t.
   Suppose $\kappa, w \models Kiq$. That is, for every complete consistent superset w', if $(w, w') \in Ri$ then $q \in w'$. Now, we prove $Kiq \in w$. If $Kiq \notin w$ then, according to Lemma 5.2, there must be a consistent complete superset w', such that $(w, w') \in Ri$ and $\neg q \in w'$. This is a contradiction, so $Kiq \in w$.
   On the other hand, suppose $Kip \in w$. For every w', if $(w, w') \in Ri$, then it is obvious that $p \in w'$. So, according to the induction step, $\kappa, w' \models p$. Hence, $\kappa, w \models Kiq$. That is $\kappa, w \models p$

Notice w0 is a consistent complete superset of T, $T \subseteq w0$. From proposition 5.3, 5.4, we can get theorem 5.5

Theorem 5.5 *Suppose T is a consistent theory. Then, for every Canonical W-Kripke structure of T,* $\kappa = (W, \sigma, w0, R0, ..., Rn)$ *is a W-Kripke model of T.* □

Lemma 5.3 If $p \notin Cons(T)$, then we can choose an initial world w0 for the Canonical W-Kripke structure of T, such that $\kappa, w0 \models \neg p$ □

Suppose theory T is a consistent theory, Can(T) is the set of all the Canonical W-Kripke structures of T. Obviously, it is a subset of the models of T. From Lemma 5.3, it is easy to see that $Cons(T) = \{p | p$ is valid in all structures of $Can(T)\}$. So, we get the following theorem:

Theorem 5.6 *Completeness of W.*

Suppose T is a consistent theory. Then, all the semantic entailment of T is the consequence conclusion of T. In other words, $Th(T) \subseteq Cons(T)$. □

Theorem 5.7 *Complete Theorem*

1. *Formula p is consistent iff p is satisfiable.*
2. *For every consistency theory T,* $Th(T) = Cons(T)$. □

6 Conclusion

There are great differences between logic system W and preciously developed knowledge or belief logic systems based on a multi-agent reasoning system. The main differences are: 1. W does not contain the necessary rule. 2. W allows an normal agent's knowledge to be inconsistent with the real world knowledge. 3. W is specially designed for multi-agent systems based on shared common knowledge views. So, it provides complete axioms about common knowledge's properties and how to use common knowledge. There are also many works on W, including its abilities, relationship with other knowledge systems, and algorithm about its consistency etc. Related research results will be published in the future [26] [27].

Acknowledgments

The first author would like to thank Professor Kazuhiro Fuchi, pre-director of the ICOT research center, Dr. Shunichi Uchida, the director of the ICOT research center, Professor Koichi Furukawa, pre-vice-director of the ICOT research center for their encouragement. Special thanks go to Dr. Katsumi Nitta, manager of the second research laboratory, Dr. Akira Aiba and all the colleagues of the second research laboratory of ICOT for their discussion and valuable suggestions. Great gratitude goes to Mr. K. Narita for his great help while the first author worked and lived in Japan.

References

[1] Barwise, J., Three Views of Common Knowledge, TARK'88, pp.365-380.

[2] B.F. Chellas, Modal Logic, Cambridge University Press, 1980.

[3] Fagin, R. and Halpern, J.Y., Reasoning about Knowledge and Probability: Preliminary Report, TARK'88, pp.277-293.

[4] Fagin, R., and Joseph Y. Halpern, Belief, Awareness, and Limited Reasoning: Preminary Report, IJCAI'85, pp491-501.

[5] Fagin, R., Halpern, J.Y., and Vardi M.Y., A Model-Theoretic Analysis of Knowledge, JACM, Vol. 38, No. 2, April 1991, pp. 382-428.

[6] Hapern, J.Y. and Moses, Y., Knowledge and Common Knowledge in Distributed Environment, in Proc. 3rd. ACM Symp. on Principles of Distributed Computing, pp50-61. Revised and expanded version of this paper with the same title later appeared on JACM, Vol. 37, NO. 3, July 1990, pp. 549-587.

[7] Halpern, J.Y. and Moses, Y., A Guide to the Modal Logic of Knowledge and Belief, IJCAI'85, pp.480-490.

[8] J. Hintikka, Impossible Possible Worlds Vindicated, J. Philosophical Logic 4 (1975), pp.475-484.

[9] J. Hintikka, Knowledge and Belief, Cornell University Press, 1982.

[10] Konolige,K., Circumscriptive Ignorance, AAAI-82, pp202-204.

[11] Konolige, K., A Computational Theory of Belief Introspection, IJCAI'85, pp502-508.

[12] Lewis, D., Convention, A Philosophical Study, 1969, Harvard U. Press.

[13] D. Lehmann, Knowledge, Common Knowledge, and Related Puzzles, in Proceedings of the 3rd Annual ACM Conference on Principle of Distributed Computing, 1984, pp 62-67.

[14] Levesque, A Logic of Implicit and Explicit Belief, in Proceedings of the 1984 National Conference on AI, AAAI-84, pp198-202.

[15] J. McCarthy, M. Sato, T. Hayashi, S. Igarashi, On the Modal Theory of Knowledge, Computer Science Technological Report STAN-CS-78-657, Stanford University, April 1978.

[16] Marek, V.W. and Truszczynski, M., Non-monotonic Logic, Context-Dependent Reasoning, Springer-Verlag, 1993.

[17] Nakashima, H., Peter, S., and Schutze, H., Communication and Inference through Situations, IJCAI'91, pp.76-81.

[18] Schiffer, S., Meaning, 1972, Oxford University Press.

[19] Yoav Shoham, Yoram Moses, Belief as Defeasible Knowledge, IJCAI'89, 1168-1172. An extented and completed version with the same title has been published on Journal of Artificial Intelligence 64 (1993) 299-321.

[20] Vardi, M.Y. A Model-Theoretic Analysis of Monotonic Knowledge, IJCAI'85, pp509-512.

[21] Vardi (ed.), Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge, 1988.

[22] Wang Xianchang et al., W - A Logic System Based on Shared Common Knowledge Views. In the Proceedings of IJCAI'93, pp410-415.

[23] Wang Xianchang and Chen Huowang, On Assumption Reasoning in Multi-Reasoner System, In the Proceedings of Pacific Rim International Conference on Artificial Intelligence, 1990.

[24] Wang Xianchang, Non-Monotonic Reasoning and Non-Monotonic Reasoning System -WMJ. Ph.D. Thesis, 1991, Changsha Institute of Technology, P.R. China.

[25] Wang Xianchang, Knowledge and Common Knowledge : Their Representation and Reasoning, Journal of Computer Science, (in Chinese), vol. 1, pp9-16, 1993.

[26] Wang Xianchang, Fool's Logic: The Shared Common Knowledge Multi-Agent System's Model, Institute for New Generation Computer Technology (ICOT), 1994.

[27] Wang Xianchang, Relationship Between Multi-Agent Logic System W and Weak S5 Systems, Institute for New Generation Computer Technology (ICOT), 1994.