UDC: 007.52
Original scientific paper

ON THE NOTION OF *BEING MORE KNOWLEDGEABLE* IN MULTI-AGENT SYSTEMS

Mirko Čubrilo
University of Zagreb, Faculty of Organization and Informatics, Varaždin, Croatia
E-mail:mcubrilo@foi.hr

Mirko Maleković
University of Zagreb, Faculty of Organization and Informatics, Varaždin, Croatia
E-mail:mmalekov@foi.hr

In this paper we shall consider the notion of being more knowledgeable *in multi-agent systems. We shall demonstrate that the relatively simple relationship between the possibility relations for two agents, which generalizes the inclusion relationship, possesses some of the intuitive features we expect from such a notion.*

Keywords: knowledge operators, multi-agent systems, the notion of *being more knowledgeable* in multi-agent systems.

1. INTRODUCTION

The theory of multi-agent systems is described in [4]. An incorporation of knowledge (knowledge operators K_i, $i = 1,.., m$) is given in [5] and the time (the past temporal operators) is given in [6] and in [2].

In this paper we shall propose a relationship between the possibility relations of two agents in multi-agent systems, naturally capturing some of the properties of the notion of *being more knowledgeable* in such systems. This holds for *positive* as well as *negative* knowledge (ignorance).

2. BASIC NOTIONS

In this section, we shall introduce the basic concepts and notations. Our exposition follows that of [2].

Suppose we have a group consisting of m agents, named $1, 2,.., m$. We assume that these agents wish to reason about a world that can be described in terms of a nonempty set P of primitive propositions. A language is just a set of formulae, where the set of formulae LK of interest to us is defined as follows:

(1) The primitive propositions in P are formulae.
(2) If F and G are formulae, then so are $\neg F$, $(F \wedge G)$, $(F \vee G)$, $(F \Rightarrow G)$, $(F \Leftrightarrow G)$, and $K_i(F)$ fo r all $i \in \{1, 2, .., m\}$, where K_i is a modal operator.

A Kripke structure M for an agent group $\{1, 2,.., m\}$ over P is a tuple $M = (S, I, k_1, k_2, .., k_m)$, where S is a set of possible worlds, I is an interpretation that associates with each world in S a truth assignment to the primitive propositions in P, and $k_1, k_2,.., k_m$ are binary relations on S, called the possibility relations for agents $1, 2,.., m$, respectively.

Given $p \in P$, the expression $I[w](p) = true$ means that p is true in a world w in a structure M. The fact that p is false, in a world w of a structure M, is indicated by the expression $I[w](p) = false$.

The expression $(u, v) \in k_i$ means that an agent i considers a world v possible. given his information in a world u. Since k_i defines what worlds an agent i considers possible in any given world, k_i will be called the possibility relation of the agent i.

We now define what it means for a formula to be true at a given world in a structure. Let $(M,w) \models F$ mean that F holds or is true at (M, w). The definition of $\models$ is as follows:

(a) $(M, w) \models p$ iff $I[w](p) = true$, where $p \in P$
(b) $(M, w) \models F \wedge G$ iff $(M, w) \models F$ and $(M, w) \models G$
(c) $(M, w) \models F \vee G$ iff $(M, w) \models F$ or $(M, w) \models G$
(d) $(M, w) \models F \Rightarrow G$ iff $(M, w) \models F$ implies $(M, w) \models G$
(e) $(M, w) \models F \Leftrightarrow G$ iff $(M, w) \models F \Rightarrow G$ and $(M, w) \models G \Rightarrow F$
(f) $(M, w) \models \neg F$ iff $(M,w) \not\models F$, that is, $(M, w) \models F$ does not hold
(g) $M \models F$ iff $(M, w) \models F$ for all $w \in S$

Finally, we shall define a modal operator K_i, where $K_i(F)$ is read: Agent i knows F.

(h) $(M, w) \models Ki(F)$ iff $(M, t) \models F$ for all $t \in S$ such that $(w, t) \in k_i$.

In (h) it is the case that an agent i knows F in a world w of a structure M exactly if F holds at all worlds t that the agent i considers possible in w.

2.1. Multi-agent systems

A multi-agent system is any collection of interacting agents. Our key assumption is that if we look at the system at any point in time, each of the agents is in some state. We refer to this as the agent's local state. We assume that an agent's local state encapsulates all the information to which the agent has access. Since each agent has a local state, it is very natural to think of the whole system as being in some (global) state. The global state includes the local states of the agents and the local state of an environment. Accordingly, we divide a system into two components: the environment and the agents, where we view the environment as everything else that is relevant. Also, the environment can be viewed as just another agent. We must say that a given system can be modelled in many ways. How to divide the system into agents and environment depends on the system being analyzed.

Let L_e be a set of possible local states for the environment and let L_i be a set of possible local states for agent i, $i = 1, .., n$. We define $G = L_e \times L_1 \times .. \times L_n$ to be a set of global states. A global state describes the system at a given point in time. What we are interested in is how the systems (since they are not static entities) change over time. We take time to range over the natural numbers, that is, the time domain is the set of natural numbers, N.

A run over G is a function $r : N \rightarrow G$.

Thus, a run over G can be identified with a sequence of global states in G. The run r represents a complete description of how the system's global state evolves over time. $r(0)$ describes the initial global state of the system in a possible execution, $r(1)$ describes the next global state, and so on.

If $r(m) = (s_e, s_1,.., s_n)$, then we define $r[e](m) = s_e$ and $r[i](m) = s_i$, for $i=1,..,n$.

Note that $r[i](m) = s_i$ is the local state of the agent i at the (global) state $r(m)$.

A system R over G is a set of runs over G. The system R models the possible behaviours of the system being modelled. The intuitive understanding that the system being modelled has certain patterns of behaviour can be captured by the requirement that the set of runs R is nonempty. At the end of this section, let's highligt the fact that the authors in [4] showed how to model a knowledge base as a multi-agent system. It is a very interesting result because in doing so we can talk about what the knowledge base knows with regard to its knowledge. Some problems in this approach will be discussed in a forthcoming paper.

2.2. Knowledge in multi-agent systems

We assume that we have a set P of primitive propositions, which we can think of as describing basic facts about a system R. Let I be an interpretation for the propositions in P over G, which assigns truth values to the primitive propositions at the global states. Thus, for every $p \in P$ and $s \in G$, $I[s](p) \in \{true, false\}$. An interpreted system IS is a pair (R, I).

Now, we define knowledge in an interpreted system IS.

Let $IS = (R, I)$ be an interpreted system. A Kripke structure for IS, denoted by $M(IS) = (S, I, k_1, .., k_n)$, is defined in a straightforward way.

$S = \{r(m) \mid r \in R, m \in N\}$, that is, S is the set of the global states at the points (r, m) in the system R.

The possibility relations $k_1, k_2,.., k_n$ are defined as follows:

Let $r(m) = (s_e, s_1,.., s_n)$, $r'(m') = (s_e', s_1',.., s_n')$ be global states in S. We say that $r(m)$ and $r'(m')$ are indistinguishable to an agent i iff $s_i = s_i'$.

Thus, the agent i has the same local state in both $r(m)$ and $r'(m')$. We define $k_i = \{(r(m), r'(m')) \in S \times S \mid r(m)$ and $r'(m')$ are indistinguishable to the agent $i\}$, $i=1, 2,.., n$.

Accordingly, $(r(m), r'(m')) \in k_i$ iff $s_i = s_i'$, $i = 1, 2,.., n$.

There is no possibility relation k_e for the environment because we are not usually interested in what the environment knows.

Now it is evident what exactly it means for a formula F in LK to be true at a state $r(m)$ in an interpreted system IS. For instance, we have

$(IS, r(m)) \models p$ iff $I[r(m)](p) = true$, for all $p \in P$.

$(IS, r(m)) \models K_i(F)$ iff $(IS, r'(m')) \models F$ for all $r'(m') \in S$ such that $(r(m), r'(m')) \in k_i$.

We say that the formula F in LK is valid in an interpreted system IS, and is denoted by $IS \models F$, iff $(IS, r(m)) \models F$ for all $r(m) \in S$.

To be able to make temporal statements, we extend our language LK by adding temporal operators, which are new modal operators for talking about time. The language (LK + "The past temporal operators") will be denoted by $LKPT$, and will be used for reasoning about events that happen along a single run r (in the past) in the system R.

The past temporal operators

A past formula (it must include at least one past temporal operator) describes a property of a prefix of the state, lying to the left of the current position, that is, a past formula at the state $r(m)$ describes a property of the states $r(0), r(1),..., r(m-1), r(m)$.

The Previous Operator po If $F \in LKPT$, then so is poF, read as F was previously. Its semantics is defined by $(IS, r(m)) \models poF$ iff $m > 0$ and $(IS, r(m-1)) \models F$.

Thus, poF holds at the state $r(m)$ iff $r(m)$ is not the first state in the run r and if F holds at the state $r(m-1)$. In particular, poF is false at the state $r(0)$. This operator makes sense because our notion of time is discrete. All the other past temporal operators make perfect sense even for continuous notions of time.

The Has-always-been Operator $p\heartsuit$ $p\heartsuit F \in LKPT$ if $F \in LKT$. It is read has always been F, and is defined by $(IS, r(m)) \models p\heartsuit F$ iff (for all m', $0 \le m' \le m)[(IS, r(m')) \models F]$

Thus, $p\heartsuit F$ holds at state $r(m)$ iff F holds at the state $r(m)$ and all the preceding positions.

The Once Operator $p\diamond$ If $F \in LKPT$, then so is $p\diamond F$, read once F. Its semantics is defined by $(IS, r(m)) \models p\diamond F$ iff (for some m', $0 \le m' \le m)[(IS, r(m')) \models F]$.

Accordingly, $p\diamond F$ holds at the state $r(m)$ iff F holds at the state $r(m)$ or some preceding state.

The Sometime-has-not-been Operator $\nrightarrow$ If $F \in LKPT$, then so is $\nrightarrow F$, read *sometime-has-not-been F*. Its semantics is defined by $(IS, r(m)) \models \nrightarrow F$ iff (for some m', $0 \leq m' \leq m)[(IS, r(m')) \models \neg F$].

Thus, $\nrightarrow Ki(F)$ holds at the state $r(m)$ iff there was a time point in the past (say m'), such that $(IS, r(m')) \nvDash F$.

The possibility relations are usually equivalence relations (reflexive, symetric and transitive). They can also have a few additional features, but here we do not assume *any* of the special features of the possibility relations.

3. THE NOTION OF BEING MORE KNOWLEDGEABLE IN MULTI-AGENT SYSTEMS

In everyday life we are used to saying that a person A is *more knowledgeable* (has greater knowledge) than a person B. In this context this notion does not exclusively signify *strictly more*, but it also embraces the equality of knowledge. After all, the meaning of this notion is primarily intuitive for us and we don't usually try to analyze it. But, in multi-agent systems, and especially in those systems with the cooperation structures employed (see [3]), it has to be done. The more knowledgeable an agent is, the more important it should be for the agent community as a whole.

The first candidate relationship between the possibility relations of agents A and B, for *A being more knowledgeable than B*, is of course the inclusion relationship between appropriate relations, i.e. $k_B \subseteq k_A$. But we can generalize this relationship (but probably not in a unique manner) in the sense that the new relationship captures all the features of the old one, at least while considering the notion of *being more knowledgeable*.

As our candidate relationship between the possibility relations of the agents i and j to capture the notion of *being more knowledgeable* in multi-agent systems we choose a relationship which has emerged in the investigations concerning error recovery in databases [1]. It is defined in Definition 3.

Definition 1: Relation $R^{\vee}$ on a set X is an *inverse relation* of the relation R on that set if $R^{\vee}=\{(y,x):(x,y)\in R\}$.

Definition 2: Relation R on the set X is *more deterministic* than the relation S (denoted $R \succcurlyeq S$) on the same set, provided that $R^{\vee}\circ R \subseteq S^{\vee}\circ S$

Proposition 1: If $R \subseteq S$ then $R \succcurlyeq S$

Proof: Obvious

The following small example demonstrates that the converse proposition of Proposition 1 is generally untrue.

Let $R = \{(a,b),(a,c),(x,y)\}$, $S = \{(a,b),(a,c),(y,y)\}$.

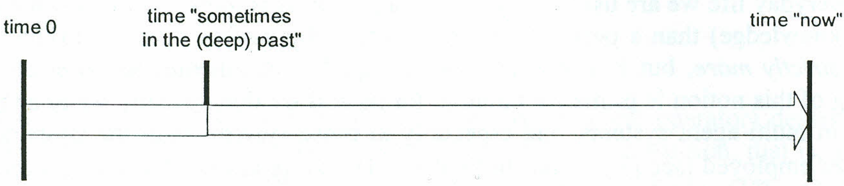
Then $R^{\vee} = \{(b,a),(c,a),(y,x)\}$ and $S^{\vee}=\{(b,a),(c,a),(y,y)\}$.

Further, we have $R^{\vee}\circ R = \{(b,b), (b,c), (c,b), (c,c), (y,y)\}$ and $S^{\vee}\circ S = \{(a,a), (b,b), (a,b), (b,a), (y,y)\}$.

Now it is clear that $R^{\vee}\circ R \subseteq S^{\vee}\circ S$ holds, but $R \subseteq S$ does not.

Definition 3: If $k_j \succcurlyeq k_i$, with k_i and k_j representing the respective possibility relations of the agent *i* and *j*, we say that agent *j* is *more knowledgeable* than agent *i*.

In everyday life when we say that person A knows more (is more knowledgeable) than person B, we do not only mean that person A's knowledge is greater than person B's knowledge at that very moment, but (at least implicitly) that this was true during the whole past or at least a part of the past (see Picture 1).



Picture 1

The most important past temporal operator is the *has-always-been* [6]. The following proposition captures the implicit supposition we have just mentioned; this is the one concerning the notion of a person A being more knowledgeable than a person B in multi-agent systems.

Proposition 2: If *has-always-been* that the agent *i* knew *F* and the possibility relation of the agent *j* is more deterministic than the same relation of the agent *i* (i.e., if the agent *j* is more knowledgeable than the agent *i*) then it *has-always-been* that the agent *j* knew *F*. More formally:

$$\text{If } k_j \succcurlyeq k_i \text{ and } (IS,\ r(m)) \models p\,\heartsuit K_i(F), \text{ than } (IS,\ r(m)) \models p\,\heartsuit K_j(F).$$

Proof:

$$(IS,\ r(m)) \models p\,\heartsuit K_i(F) \equiv$$
$$\equiv (\forall m_i,\ 0 \le m_i \le m)((IS,\ r(m_i)) \models K_i(F)) \equiv$$
$$\equiv (\forall m_i,\ 0 \le m_i \le m)(\forall m'_i,\ (r(m_i),r(m'_i)) \in k_i)((IS,\ r(m'_i)) \models F) \quad (1)$$

Now, suppose that $(IS,\ r(m)) \not\models p\,\heartsuit K_j(F)$. We then know that

$$(\exists m_j,\ 0 \le m_j \le m)(\exists m'_j,\ (r(m_j),r(m'_j)) \in k_j)((IS,\ r(m'_j)) \not\models F) \quad (2)$$

Now, (1) and 2 $\Rightarrow (r(m_j),r(m'_j)) \notin k_i \Rightarrow (r(m'_j),r(m_j)) \notin k_i^{\vee} \Rightarrow (r(m'_j),r(m'_j))\ k_i^{\vee}\circ k_i$ (3)

On the other hand, from (2) we get $(r(m'_j),r(m'_j)) \in k_j^{\vee}\circ k_j$ (4)

Finally, (3) and (4) both contradict the Proposition hypothesis that $k_j \succcurlyeq k_i$.

Proposition 3: If an agent *i* knows formula F and the possibility relation of an agent *j* is more deterministic than that of an agent *i* ($k_j \succcurlyeq k_i$), then the agent *j* also knows *F*. More formally:

$$\text{If } (IS, r(m)) \models K_i(F) \text{ and } k_j \succcurlyeq k_i, \text{ then } (IS, r(m)) \models K_j(F)$$

Proof: Analogous with the proof of Proposition 2.

Proposition 3 is our *minimal* proposition concernig the relation $\succcurlyeq$. It is a *must* for $\succcurlyeq$ to be the right candidate relation in order to embrace the notion of *being more knowledgeable*. Of course, Proposition 2 generalizes Proposition 3. It is much stronger than Proposition 3 and is also very intuitive. We can extend Proposition 3 in an obvious way in order to make it hold for the remaining past temporal operators: *once*, *previous* and *sometime-has-not-been*. We now have the following proposition:

Proposition 4: If k_i and k_j are the respective possibility relations of the agents *i* and *j* and $k_j \succcurlyeq k_i$, then the following holds:

a) If $(IS, r(m)) \models p\text{-}oK_i(F)$, then $(IS, r(m)) \models p\text{-}oK_j(F)$

b) If $(IS, r(m)) \models p \blacklozenge K_i(F)$, then $(IS, r(m)) \models p \blacklozenge K_j(F)$

c) If $(IS, r(m)) \models \nrightarrow K_i(F)$, then $(IS, r(m)) \models \nrightarrow K_j(F)$

Proof: Analogous to the proof of Proposition 1.

The propositions 2 to 4 concern agents knowing something "positively". It is intuitively understandable that if agent *j* does not know something and it is more knowledgeable than agent *i*, then agent *i* does not know it either. We shall prove this in the following proposition.

Proposition 5: If an agent *j* doesn't know F and $k_j \succcurlyeq k_i$, then agent *i* doesn't know F either.

More formally:

If $(IS, r(m)) \models \neg K_j(F)$ and $k_j \succcurlyeq k_i$, then $(IS, r(m)) \models \neg K_i(F)$

Proof:

$$(IS, r(m)) \models K_j(F) \equiv (\forall m_j, (r(m),r(m_j)) \in k_j)((IS, r(m_j)) \models F) \quad (1)$$

By negating (1) we get

$$(IS, r(m)) \models \neg K_j(F) \equiv \neg((\forall m_j, (r(m),r(m_j)) \in k_j)((IS, r(m_j)) \models F)) \equiv$$
$$\equiv (\exists m_j, (r(m),r(m_j)) \in k_j)((IS, r(m_j)) \models \neg F) \quad (2)$$

Now suppose that the Proposition hypothesis that states $(IS, r(m)) \models \neg K_i(F)$ does not hold.

$$(IS, r(m)) \not\models \neg K_i(F) \equiv (IS, r(m)) \models K_i(F) \equiv$$
$$\equiv (\forall m_i, (r(m),r(m_i)) \in k_i)((IS, r(m_i)) \models F) \quad (3)$$

From (2) and (3) we conclude that $(r(m),r(m_j)) \notin k_i$, and that eventually,

$$(r(m_j),r(m_j)) \notin k_i^{\vee} \circ k_i \quad (4)$$

From (3) we eventually conclude that $(r(m_j),r(m_j)) \in k_j^{\smile} \circ k_j$ (5). Now, (4), (5) and the Proposition hypothesis which states $k_j \not\succeq k_i$ lead to a contradiction.

4. CONCLUSIONS

Our candidate relationship between the possibility relations of two agents, intended to capture the notion of being more knowledgeable in multi-agent systems, turns out to be good at capturing some of the important intuitive features of the notion that is being considered. Of course, one has to investigate more in order to convince oneself that the chosen formalization of the notion of being more knowledgeable is the right one.

REFERENCES

[1] C. Brink, W. Kahl and G. Schmidt (eds). *Relational Methods in Computer Science*. Springer, Berlin, 1997.

[2] M. Čubrilo and M. Maleković. Some properties of *sometime-has-not-been* temporal operator in multi-agent systems. *ITI '99*, Pula, Croatia, 1999.

[3] M. d'Inverno, M. Luck and M. Wooldrige. Cooperation structures. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence*, Nagoya, Japan, 1997, pp. 600 - 605.

[4] R. Fagin, J. Y. Halpern, Y. Moses and M. Y. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, Mass., 1995.

[5] M. Maleković. Integrated knowledge. *Proc. of the 20th Int. Conf. ITI '98*, Pula, 16-19 July 1998, ed. by A. Kalpić and V. H. Dobrić, pp. 93-100.

[6] M. Maleković. Agent properties in multi-agent systems (accepted for publishing in *Informatica: an International Journal of Computing and Informatics*).

[7] J. van Benthem. *Modal Logic and Classical Logic*. Bibliopolis, Napoli, 1983.

Received: 12 October 1999
Accepted: 17 December 1999

Mirko Čubrilo
Mirko Maleković

O POJMU "ZNATI VIŠE OD" (DRUGIH AGENATA) U VIŠEAGENTNIM SUSTAVIMA

Sažetak

U radu se analizira značenje pojma "znati više" u višeagentnim sustavima. Pojam se formalizira kroz odnos (metarelaciju) na relacijama zora pojedinačnih agenata višeagentnih sustava. Pokazuje se da jedna relacija, ranije proučavana u kontekstu teorije uklanjanja grešaka u bazama podataka, ima prirodna svojstva relacije "znati više". Daju se i dokazuju propozicije koje pokazuju da je izabrana relacija sukladna intuitivnom i formalnom poimanju temporalnih operatora i operatora znanja.

Ključne riječi: višeagentni sustavi, relacija zora, (meta)relacija "znati više", temporalni operatori, operatori znanja.