Beyond Public Announcement Logic
An Alternative Approach to some AI Puzzles

Paweł Garbacz, Piotr Kulicki, Marek Lechniak, and Robert Trypuz

John Paul II Catholic University of Lublin,
al. Racławickie 14, 20-950 Lublin, Poland
{garbacz,kulicki,lechniak,trypuz}@kul.pl
http://www.l3g.pl

Abstract. In the paper we present a dynamic model of knowledge. The model is inspired by public announcement logic and an approach to a puzzle concerning knowledge and communication using that logic. The model, using notions of situation and epistemic state as foundations, generalizes structures usually used as a semantics for epistemic logics in static and dynamic aspects. A computer program automatically solving the considered puzzle, implementing the model, is built.

Key words: knowledge representation, dynamic epistemic logic, multi-agent systems, Prolog

1 Introduction

Knowledge representation plays an important role in AI. Moore in [7] claims that the basic application of logic in AI is an analysis of this concept. He points out two uses of formal logic in that context. Firstly, the formalism of logic can be used as a knowledge representation system; secondly, logic provides forms of deductive reasoning. The formalism of logic has two components: syntactic (language and axioms) and semantic (model theoretic structure). In the case of knowledge representation the language of logic includes epistemic operators that enable us to formulate sentences of the type: "agent a knows that φ","an agent a believes that φ", etc. Axioms set up a theory that defines valid deductions for such sentences. The semantic component of epistemic logic is usually based on structures built from sets of possible worlds. In that setting logical space of epistemically possible worlds, i.e. worlds consistent with an agent's knowledge, is considered.

McCarthy [6] argues, that from the AI point of view the semantic part of epistemic logic is more useful. The reason is that it is more intuitive and easier from the operational perspective. We share that opinion. Moreover, we think that philosophical understanding of semantic structures for epistemic logic still requires more attention. Thus, our main effort is to construct an ontologically founded model for epistemic logic. In that model, we would like to be able to represent a vast spectrum of notions about knowledge in multi-agent environments, including not only the actual knowledge of the agent himself and other agents but possible knowledge, change of knowledge and relations between knowledge and actions as well.

As a working example we will use a well-known *hat puzzle: Three people Adam, Ben and Clark sit in a row in such a way that Adam can see Ben and Clark, Ben can see Clark and Clark cannot see anybody. They are shown five hats, three of which are red and two are black. The light goes off and each of them receives one of the hats on his head. When the light is back on they are asked whether they know what the colours of their hats are. Adam answers that he doesn't know. Then Ben answers that he doesn't know either. Finally Clark says that he knows the colour of his hat. What colour is Clark's hat?*

2 Solving puzzles in Public Announcement Logics

Public announcement logic [8, 2, 10] (henceforward PAL) is an extension of epistemic logic [11, 4], which is a modal logic formally specifying the meaning of the formula $K_a\varphi$—"agent a knows that φ". PAL enriches epistemic logic with a dynamic operator of the form "$[\varphi]\psi$", which is read "after public and truthful announcement that φ, it is the case that ψ". That extension allows for reasoning about the change of agent's beliefs being a result of receiving new information.

The language of PAL is defined in Backus-Naur notation as follows:

$$\varphi \ ::= \ p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_a\varphi \mid [\varphi]\varphi$$

where p belongs to a set of atomic propositions Atm, a belongs to a set of agents $Agent$ and K is an operator of individual knowledge and $[\_]$ is an operator of public announcement.

Models of PAL are Kripke structures with valuation function and have the form $\mathcal{M} = \langle S, Agent, \sim, v\rangle$. Its components are described as below.

- S is a set of situations (or possible worlds).
- $Agent$ is a set of rational agents.
- $\sim$ is a function assigning a set of pairs of situations about which it is said that they are *indiscernible* to every element of $Agent$. Thus $s \sim_a s'$ means that for agent a, two situations, s and s', are indiscernible.
- $v : Atm \longrightarrow 2^S$ is a standard valuation function.

Satisfaction conditions for PAL in the model $\mathcal{M}$ and for $s \in S$ are defined in a standard way. For knowledge and public announcement operators they have the forms:

$$\mathcal{M}, s \models K_a\varphi \iff \forall s' \in S\ (s \sim_a s' \Longrightarrow \mathcal{M}, s' \models \varphi)$$
$$\mathcal{M}, s \models [\varphi]\psi \iff \mathcal{M}, s \models \varphi \Longrightarrow \mathcal{M}^\varphi, s \models \psi$$

where $\mathcal{M}^\varphi = \langle S', A, \sim', v'\rangle$ is characterized in such a way that: $S' = \{s \in S : \mathcal{M}, s \models \varphi\}$, for every $a \in Agent$, $\sim'_a = \sim_a \cap (S' \times S')$ and for every $p \in Atm$, $v'(p) = v(p) \cap S'$.

The epistemic logic S5 is sound and complete with respect to $\mathcal{M}$, when $\sim_a$ are equivalence relations. One can find the axiomatisation of "[_]" in [10, section 4].

The concept of knowledge presupposed by PAL is such that agent knows that φ when in all situations (possible words) belonging to his epistemic possibility space, φ holds. The epistemic possibility space of agent a can be seen as a set of situations which are equally real for that agent.

By reducing the epistemic possibility space the agent minimizes his level of ignorance [5]. An omniscient agent would have in his epistemic possibility space just one single world, i.e. the real one, whereas an agent having no knowledge would consider all situations possible, or in other words, equally real.

The reduction of the level of ignorance is done in PAL by a public and truthful announcement about a real situation. For instance, an agent having no knowledge, just after receiving the information that he is wearing a red hat, will restrict his epistemic possibility space to those situations in which it is true that he is wearing a red hat.

Now let us return to the hats puzzle. The model for the hats puzzle, $\mathcal{M}^* = \langle S, Agent, \sim, v\rangle$, is such that $Agent = \{Adam, Ben, Clark\}$. Remaining elements of $\mathcal{M}^*$ are described below.

In the models of PAL a situation is usually represented by a tuple. For instance $\langle r, r, w\rangle$ is 3-tuple representing the situation in which Adam has a red hat (r), Ben has a red hat, too (r) and Clark has a white hat (w). Thus:

$$S = \{\langle r,r,r\rangle, \langle r,r,w\rangle, \langle r,w,r\rangle\ \langle r,w,w\rangle, \langle w,r,r\rangle, \langle w,r,w\rangle, \langle w,w,r\rangle\}.$$

Because the indistinguishability relation (relativized to agent a), $\sim_a$, is an equivalence relation, the equivalent classes and a quotient set (partition), $S/_{\sim_a}$, can be defined in a standard way. Intuitively, each cell of this partition is a set of the situations which are indistinguishable for agent a.

Which situations are indistinguishable for each of the three men? It is a question on which PAL can answer neither on language nor on semantic level. Although in the hats puzzle, before the men say anything, an indistinguishability relation can be precisely defined by means of *sees* relation. $sees(a, b)$ is an irreflexive relation (because it is assumed in the puzzle that nobody sees himself) expressing a fact that a sees (perceives) b. Let $\alpha^i_{a_i}$ be an expression saying that agent a_i has a hat of the colour α^i. Then:

$$\langle \alpha^1_{a_1}, \alpha^2_{a_2}, \alpha^3_{a_3}\rangle \sim_{a_i} \langle \beta^1_{a_1}, \beta^2_{a_2}, \beta^3_{a_3}\rangle \iff \forall j(sees(a_i, a_j) \to \alpha^j = \beta^j), \quad (1)$$

where $1 \leqslant i, j \leqslant 3$ and $\alpha^1, \ldots, \beta^3$ represent red or white colour and a_1, a_2, a_3 represent Adam, Ben and Clark respectively.

Because Adam sees the (colours of the) hats of Ben and Clark, and he does not see his own hat, the epistemic possibility space for Adam is described by the set:

$$S/_{\sim_{Adam}} =$$

$$\{\{\langle r,w,r\rangle, \langle w,w,r\rangle\}, \{\langle r,r,r\rangle, \langle w,r,r\rangle\}, \{\langle r,r,w\rangle, \langle w,r,w\rangle\}, \{\langle r,w,w\rangle\}\}$$

Similarly we get the sets $S/_{\sim_{Ben}}$ and $S/_{\sim_{Clark}}$:

$$S/_{\sim_{Ben}} = \{\{\langle r,w,r\rangle, \langle w,w,r\rangle, \langle r,r,r\rangle, \langle w,r,r\rangle\}, \{\langle r,r,w\rangle, \langle w,r,w\rangle, \langle r,w,w\rangle\}\}$$

$$S/_{\sim_{Clark}} = \{\{\langle r,r,r\rangle, \langle w,r,r\rangle, \langle r,w,r\rangle, \langle w,w,r\rangle, \langle r,r,w\rangle, \langle w,r,w\rangle, \langle r,w,w\rangle\}\}$$

Now we are going to show how PAL can deal with the hats puzzle. Let us start with the assumption that set Atm has the following elements:

$$Atm = \{has(Adam,r), has(Adam,w), has(Ben,r), has(Ben,w), has(Clark,r), has(Clark,w)\} \tag{2}$$

where $has(a,r)$ and $has(a,w)$ mean that a has a red hat and a has a white hat respectively.

Because in the puzzle each man is allowed only to announce that he knows or does not know the colour of his hat, it is useful to introduce the following definition:

$$K_a Hatcolour =_{df} K_a has(a,r) \vee K_a has(a,w) \tag{3}$$

$K_a Hatcolour$ means that a knows the colour of his hat.

Now we are ready to express in PAL what has happened in the puzzle's scenario, i.e. that $Adam$ and then Ben announce that they do not know the colours of their hats, what finally leads to the situation in which $Clark$ knows that his hat is red. Formally:

$$[\neg K_{Adam} Hatcolour][\neg K_{Ben} Hatcolour] K_{Clark} has(Clark,r) \tag{4}$$

Clark gets the knowledge about the colour of his hat by reducing his set of indiscernible situations to the set in which it is true that neither Adam knows the colour of his hat nor does Ben. As a result he receives the following set of possible situations: $\{\langle r,w,r\rangle, \langle w,w,r\rangle, \langle r,r,r\rangle, \langle w,r,r\rangle\}$ and therefore formula 4 is true. Clark taking into account what $Adam$ and Ben have said is still unable to identify the real situation. However in all remained indiscernible situations $Clark$ can be sure that he has a red hat.

In the solution of the puzzle we used both PAL and its model. This points out a more important property of PAL, namely that in order to completely understand what knowledge is in this approach we need both the language of PAL and its model. In the language of PAL we could express that an agent knows something and state by axioms of S5 some of the formal properties of the agent's knowledge. The models of PAL give us the understanding of the fact that agent a knows something in situation s by reference to the indiscernibility relation.

3 Towards a general model of epistemic change

Our model of epistemic change has two components: ontological and epistemological. The ontological part represents, in rather rough and ready way, the world our knowledge concerns. The epistemological part of the model represents the phenomenon of knowledge in its static and dynamic aspects. The situation at stake may have any ontic structure. Thus, there are situations "in which" certain objects possess certain properties, situations "in which" certain objects participate in certain relations or processes, etc.

A situation is *elementary* if no other situation is part of it, eg. *that Adam has a red hat* would be an elementary situation and *that both Adam and Ben have red hats* would not be an elementary situation. Let $ElemSit$ be a set of elementary ontic (possible) situations.

In the set $ElemSit$ we define the relation of copossibility ($\|$). Intuitively, $x \parallel y$ means that situation x may (ontologically) cooccur with situation y. For example, *that Adam has a red hat* is copossible with *that Ben has a red hat*, but is not copossible with *that Adam has a white hat*. The relation $\|$ is reflexive and symmetric in $ElemSit$, but is not transitive.

In what follows we will represent situations as sets of elementary situations. Let $\emptyset \notin Sit \subseteq \wp(ElemSit)$ be a set of (possible) ontic situations. Given our understanding of the relation $\|$, the following condition is accepted:

$$X \in Sit \rightarrow \forall y, z \in X y \parallel z. \tag{5}$$

We can now define the notion of a possible world:

$$X \in PossWorld \triangleq X \in Sit \land \forall Y (X \subset Y \rightarrow Y \notin Sit). \tag{6}$$

Let $Time = (t_1, t_2, \dots)$ be a sequence of moments. The *actual epistemic state* of an agent at a given moment will be represented by a subset of $PossWorld$: $epist(a, t_n) \subseteq PossWorld$. Any such state collectively, so to speak, represents both the agent's knowledge and his ignorance. Due to its actual epistemic state, which is represented by a set $epist(a, t_n)$, and for every $X \in Sit$, agent a may be described (at t_n) according to the following three aspects:

Definition 1. *Agent a knows at moment t_n that situation X holds (written: $K_{a,t_n}(X)$) iff $X \subseteq \bigcap epist(a, t_n)$.*

Definition 2. *Agent a knows at moment t_n that situation X does not hold (written: $\overline{K}_{a,t_n}(X)$) iff $X \cap (\bigcup epist(a, t_n)) = \emptyset$.*

Definition 3. *Agent a does not have any knowledge at moment t_n about situation X iff $\neg K_{a,t_n}(X) \land \neg \overline{K}_{a,t_n}(X)$.*

However, the puzzles we are dealing with do not presuppose that we know the actual epistemic state of a given agent. Thus, we extend the notion of actual epistemic state to the notion of possible epistemic state. A *possible epistemic state* of an agent represents a body of knowledge (resp. of ignorance) that the agent may exhibit given the ontic situation and epistemic capabilities of the agent. In our case, the possible epistemic states are determined by the relation of seeing, other agents' announcements and the agent's deductive capabilities.

A possible epistemic state of an agent at a given moment will be represented by the set $epist_i(a, t_n) \subseteq Sit$.

Definition 4. *Agent a knows in a possible epistemic state X that (ontic) situation Y holds (written :* $K_{epist_i(a,t_n)}(Y)$*) iff* $Y \subseteq \bigcap X$*.*

Definitions analogous to 2 and 3 can be also added.

We will use the following auxiliary notions:

- $Epist(a, t_n)$ - the set of all possible epistemic states of agent a at t_n,
- $Epist(t_n)$ - the set of sets of possible epistemic states of all agents at t_n,
- $Epist$ - the set of sets of all possible epistemic states of all agents (from $Agent$) at all moments (from $Time$).

When $a \in Agent$, then "$\sim_a$" will represent the relation of epistemological indiscernibility, which we treat as an equivalence relation. In a general case, the epistemological indiscernibility covers a number of epistemic constraints of agents.

In our case, the relation of epistemological indiscernibliity depends on the knowledge obtained thanks to the behaviour of some agent.

The relation of epistemological indiscernibility is, as the notion of knowledge itself, relative to time: $\sim_{a,t_n}$ is the relation of epistemological indiscernibility for agent a at time t_n.

It is an assumption of our approach that possible epistemic states coincide with the abstraction classes of the epistemological indiscernibliity relation:

$$Epist(a, t_n) = PossWorld/ \sim_{a,t_n} \tag{7}$$

We assume that all changes of epistemic states are caused by the behaviour of agents, in particular by their utterances, by means of which they expose their (current) epistemic states and not by their inference processes.

A number of rules that govern the dynamics of epistemic states can be defined. In the hats puzzle the only rule that sets the epistemic states in motion is the following one:

If agent a (says that) he does not know that X holds, in an epistemic state $epist_i(a, t_n)$ a knows that X holds, then after the aforementioned utterance this state (i.e. $epist_i(a, t_n)$) is effectively impossible, i.e. we remove its elements from all possible epistemic states of all agents. Formally,

Rule 1 *If (a says that)* $\neg K_{a,t_n}(X)$ *and* $Y \in Epist(a, t_n)$ *and* $K_Y(X)$*, then for every* $a' \in Agent$*,* $Epist(a', t_{n+1}) = \delta_0(Epist(a', t_n), Y)$*, where*

Definition 5. δ_0 *maps* $Epist \times \bigcup\bigcup Epist$ *into* $Epist$ *and satisfies the following condition:*

$$\delta_0(Epist(a, t_n), X) = \begin{cases} Epist(a, t_n)\backslash\{X\}, & \text{if } X \in Epist(a, t_n), \\ (Epist(a, t_n)\backslash\{Z\}) \cup \{Z\backslash X\} & \text{if } Z \in Epist(a, t_n) \text{ and} \\ & X \cap Z \neq \emptyset, \\ Epist(a, t_n) & \text{otherwise.} \end{cases}$$

Some other possible rules are presented in [1].

It seems that the factors that trigger the process of epistemic change are of two kinds: ontological and epistemological. The ontological condition of this rule is the fact that agent a says that he does not know that a certain ontic situation holds. The epistemological condition is his epistemic state ($Y \in Epist(a, t_n)$), in which the agent knows that this situation holds ($K_Y(X)$). We may represent the epistemological conditions of rules for epistemic changes by means of the notion of epistemic state. However, in order to account for the ontological conditions, we distinguish in the set $\wp(ElemSit)$ a subset $AgentBeh$ that collects types (here: sets) of ontic situations that are those conditions. An example of such type may be a set of situations in which agents say that they do not know what hat they have. In general, those conditions may be classified as agents' behaviours, which include also such "behaviours" as being silent (cf. the wisemen puzzle).

Let $a \in Agent$. A *rule for epistemic change* ρ is either

1. mapping $\rho : \bigcup Epist \times AgentBeh \times \bigcup Epist \rightarrow \bigcup Epist$ (this condition concerns rules with epistemological conditions) or
2. mapping $\rho : \bigcup Epist \times AgentBeh \rightarrow \bigcup Epist$ (this condition concerns rules without epistemological conditions).

It should be obvious that

1. if $\rho(X, Y, Z) = V$ and $X, Z \in Epist(t_n)$, then $V \in Epist(t_{n+1})$ (for rules with epistemological conditions),
2. if $\rho(X, Y) = V$ i $X \in Epist(t_n)$, to $V \in Epist(t_{n+1})$ (for rules without epistemological conditions).

The set of all such rules will be denoted by "*Rule*". For the sake of brevity, from now on we will consider only rules with epistemological conditions.

In order to obtain the solution to the puzzle at stake, one needs the following input data: set Sit, temporal sequence $Time = (t_n)$, set of epistemic agents $Agent$, set of sets of epistemic states of any such agent at the initial moment t_1: $Epist_1 = \{Epist(a, t_1) : a \in Agent\}$ and function $dist : Time \rightarrow \wp(ElemSit)$.

The evolution of sets of epistemic states is triggered by the ontological conditions according to the accepted rules of epistemic change. This implies that the following condition holds (For the sake of simplicity, we assume that function $dist$ does not trigger more than one rule at a time.):

$$\exists \rho \in Rule\ \exists X \in AgentBeh\ [\rho(Epist(a, t_n), X, Epist(a', t_n)) = Z \wedge dist(t_n) \cap X \neq \emptyset] \rightarrow Epist(a, t_{n+1}) = Z. \tag{8}$$

We also assume that epistemological states change only when a certain rule is triggered:

$$Epist(a, t_{n+1}) \neq Epist(a, t_n) \equiv \exists \rho \in Rule\ \exists X \in AgentBeh\ \exists Y \in \bigcup Epist\ [\rho(Epist(a, t_n), X, Y) = Epist(a, t_{n+1})]. \tag{9}$$

4 Conclusion

We have presented a dynamic model of knowledge for rational agents. The model is inspired by public announcement logics and their usual semantics in a form of Kripke structures. However, the basic notions of our model, situation and possible epistemic state of an agent, are more fundamental from a philosophical point of view and enable us to define the notion of a possible world. We have also introduced rules defining dynamics of agents' knowledge in the presence of other agents' behaviour. Public announcements can be seen as special cases of such a behaviour.

The model has been successfully tested on logic puzzles concerning knowledge. Just by changing parameters of the program one can solve the hats puzzle with any number of agents and different perception relations. Two of the special cases here are known from literature puzzles: *three wise men* and *muddy children*. From the point of view of programming it is important that the formulation of the model could be easily transferred into a program in Prolog (see the prolog source at www.l3g.pl).

References

1. Garbacz, P., Kulicki, P, Lechniak, M., Trypuz, R.: A formal model for epistemic interactions. In: Nguyen, N. T. , Katarzyniak, R., Janiak, A. (eds.) Challenges in Computational Collective Intelligence (Springer series Studies in Computational Intelligence) (2009) (forthcoming).
2. Gerbrandy, J. D., Groeneveld, W.: Reasoning about information change. Journal of Logic, Language and Information 6, 147–169 (1997)
3. Gomez-Perez, A., Corcho, O., Fernandez-Lopez, M.: Ontological Engineering. Springer-Verlag, London (2001)
4. Hintikka, J.: Knowledge and Belief: An Introduction to the Logic of The Two Notions. Cornell University Press, Ithaca, New York (1962)
5. Hintikka, J.: On the Logic of Perception. In: Models for Modalities. Reidel Publ. Comp. Dordrecht (1969)
6. McCarthy, J.: Modality, Si! Modal Logic, No! Studia Logica 59, 29–32 (1997)
7. Moore, R. C.: Logic and Representation. In: CSLI Lecture NotesNo 39. CSLI Publications, Center for the Study of Language and Information, Standford, California (1995)
8. Plaza, J. A.: Logic of public communications. In: Emrich, M. L., Pfeifer, M. S., Hadzikadic, M., and Ras, Z. W. (eds.) Proceeddings of the 4th International Symposium on Methodologies for Intelligent Systems, pp. 201–216 (1989)
9. Russell, B.: On Denoting. Mind 14, 479–493 (1905)
10. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Volume 337 of Synthese Library Series, Springer (2007)
11. von Wright, G. H.: An Essay in Modal Logic. Amsterdam: North Holland (1951)