Dariusz Surowik
University of Bialystok

LOGIC OF ALGORITHMIC KNOWLEDGE

Abstract. In this paper we consider the construction of a **LAK** system of temporal-epistemic logic which is used to formally describe algorithmic knowledge. We propose an axiom system of **LAK** and discuss the basic properties of this logic.

Keywords: time, algorithmic knowledge, epistemic logic, temporal logic.

Formal analysis of knowledge and beliefs focuses the attention of philosophers, logicians, IT specialists, economists and researchers of many other fields of science. The best known method of formalization of knowledge and beliefs is formalization of these terms in terms of modal epistemic logic. The main achievement of modal logics is the transformation of extensional languages (with the use of which sentences are expressed which logical values depend only on the values of the component sentences) to intensional languages. The modal epistemic logic was being developed for the needs of formalization of the term of the current knowledge. It turned out quite quickly that such a notion of knowledge cannot be described with the use of any interesting logic [5]. On the basis that an agent has knowledge of specific facts, it cannot be concluded that the agent knows anything else. It cannot be assumed that the agent knows even the most elementary logical consequences of what he actually knows. Even the simplest of mathematical facts cannot be recognized as certain, if the knowledge is interpreted in such a way as it is understood, for example, during a "normal" discussion. In order to be able to describe a notion of knowledge using modal epistemic logic, an idealization of the cognitive capabilities of agents must be performed. In the systems of modal epistemic logics it is assumed, that the agent knows all the tautologies and has immediate access to all the logical consequences of his knowledge. With such assumptions the agents would have to have an unlimited store of memory and unlimited computational power. Also, the time which is necessary to perform the

necessary calculations and inferences is not taken into consideration. The modal epistemic logic suffers therefore from an affliction called a logical omniscience. The problem of the logical omniscience causes the formal systems which use epistemic logic, not to perform well in modelling the real agents and the idealization of the real agents performed by the epistemic logics is too strong in order to use this formalism to describe the knowledge of the real agents.

In order for the modal epistemic logic to be regarded as a logic of knowledge, a notion of *implicit knowledge* has been introduced. The epistemic logics do not formally describe what the cognitive subject actually knows, but they describe what is indirectly represented in the information state of a given agent. In other words, they represent what logically results from his current knowledge. That which the agent currently knows is called an *explicit knowledge*.

From the point of view of theory of knowledge, explicit knowledge is certainly a more important knowledge than implicit knowledge. The implicit knowledge that a certain path that links all the cities in a region is the shortest way, is useful for a sales representative who wants to maximize his turnover. This knowledge should become, in the case of this representative, explicit knowledge in the process of making a decision as to, which path to choose if he wants to visit all the cities in the region.

Because the explicit knowledge is so important for agents' actions, continuous efforts are being undertaken to find newer and better logic systems which can be used to formally describe this type of knowledge. A number of solutions have been proposed. An overview of some of them as well as a broad discussion of this subject can be found in the works of [2], [5], [6]. Even though specific solutions are very different, the main strategy is the same: the agents' capabilities to reason are limited by some *ad hoc* postulates. In the case of using such an approach, the problem of logical omniscience can be avoided. This makes an epistemic logic construction possible, and it could be used to describe the explicit knowledge. Such a solution however, is not without side-effects, because, as a result of such an approach, weak epistemic logics are constructed. In weak epistemic logics, rejection of the logical omniscience is realised through limiting the agents' rationality. If solving the problem of the logical omniscience is achieved through weakening of the logic of knowledge, there is nothing else left that can constitute the agents' rationality. Weak systems of the epistemic logic are thus not a very good solution when modelling the knowledge of intelligent agents.

While constructing knowledge logic which has a more solid epistemic basis, we have to deal with the problems which occur when we formalise

a "normal" notion of knowledge, i.e. with the problem of logical omniscience. As contended by D. N. Ho in works by [4] and [5] the problem of logical omniscience is only a symptom of a greater general problem. The fact that logical omniscience is a problem arises from the fact that modal epistemic logic is not capable of modeling the behaviour of agents with limited knowledge. The approach, where agents with limited knowledge are considered, offers a natural solution to the problem of the logical omniscience of cognitive subjects. A rational agent may calculate all the consequences resulting from his explicit knowledge. Then, he will have logical omniscience[1] only, when he has a large enough source of knowledge. When the knowledge sources necessary to achieve logical omniscience are not available, then the agents do not know most of the consequences of their knowledge. Therefore, they are not logically omniscient.

Usually in discussions on the subject of the representation of knowledge which is changing over time, in the language of logic it is assumed that the agent on the basis of premises has automatic and immediate access to all logical consequences resulting from these premises. The time cost which the agent incurs in order to infer the logical consequences, is not taken into consideration. Such an assumption may be taken when the knowledge of the ideal agents is being described. When describing the knowledge of the real agents, we should take into consideration not only what the agent currently knows, and what he can deduce (whenever), but also that what the agent is capable to deduct, in specified conditions, in a specified time.

Let us consider the following example. During a *Gilotine*[2] quiz show a participant taking part in the finale has sixty seconds to choose a word which matches five other words. Some participants guess the correct word, some unfortunately do not. It happens that just after the time for providing the answer runs out, the participant, who put down an incorrect answer on paper, offers the correct answer. It could be said that the finalist had the right knowledge to provide the correct answer, however the specific conditions (pressure, stress, etc.) and limited time to provide a correct answer caused him to provide an incorrect answer or no answer at all. In the case of real agents we have to consider what they can know or what they can deduce in a specific time period. Instead of problems of the type:

the agent knows (currently) that φ,

we should analyse problems of this type:

if I ask about φ, then the agent is able to provide me with a correct answer in n time units.

Questions of the above type are not connected with the classical notion of knowledge, but with the notion of algorithmic knowledge. Algorithmic knowledge have been defined in many ways.

Binmore and Shin [1] defined algorithmic knowledge in the following way:

An agent's algorithmic knowledge is whatever the agent can infer using a Turing machine.

Another definition of this type of knowledge was provided by Halpern, Moses and Vardi [3]:

An agent is said to know a fact at a certain state if at that state he can compute that he knows that fact.

Our understanding of algorithmic knowledge is close to the definition provided by Ho [4]. He defines algorithmic knowledge as knowledge satisfying the following postulates:

1. The agent knows, that φ when φ is an element of the subject's current knowledge or φ (due to the logic used by the agent) infer logically from this knowledge
2. The agent has an algorithm which is used to deduce φ and is capable of choosing this algorithm in order to use it, when he makes a choice of φ deduction,
3. The calculations take up at most n time units[3].

In order to estimate the time complexities of algorithms used by agents, we need some model of time measure. To keep it simple let us assume that the semantic time is a discreet time with a starting moment. Let us assume then that a set of time moments is isomorphic to the set of natural numbers $\mathbb{N}$. In order to express temporal context in the language of constructed logic, we introduce the following temporal operators:
- $F^n\varphi$ – there will be φ after at most n time units
- $F\varphi$ – sometimes in the future φ

Additionally we will enrich the language with epistemic operators:
- $K_i\varphi$ – agent i knows φ.

Let us consider a multi-agent system, so that index i in epistemic operators represents the agent's number.

Syntax of the constructed language is following

$$\varphi ::= p \,|\, \neg\varphi \,|\, \varphi \to \psi \,|\, K_i\varphi \,|\, F^n\varphi \,|\, F\varphi$$

where: $n \in \mathbb{N}, i \in AG$ (*AG* is a set of agents).

Formula $K_i F^n K_i \varphi$ is read as: *the agent i knows that he will find out that φ in n time units.* If then the agent decides to deduce φ from his current knowledge, then he will do it after at least n time units. It is not only expected of the agent that he has at least one procedure to calculate φ, but also that he has the skill to choose a correct procedure leading to obtain φ in a specified time period.

It sometimes happens that an agent may deduce φ, but it is not possible to estimate time, which is needed to make necessary operations leading to obtain φ. This happens in cases when the complexity of the algorithm applied by the agent is not known or when his strategy of actions is not known. Then the following statement is considered: *there is a certain number n (we do not know its value) such that the agent is able to deduce a fact φ in at least n time units.* This statement can be formally written in our language in the following way: $K_i F K_i \varphi$.

Formula $K_i F K_i \varphi$ is weaker than formula $K_i F^n K_i \varphi$ because it says nothing about how much time the agent needs to calculate φ. The formula $K_i F K_i \varphi$ says only that calculations end after a finite time.

Now, we will give the set of axioms of the **LAK**.

Axioms of LAK

1. All propositional tautologies of the language $\mathfrak{L}_{\textbf{LAK}}$.

2. Temporal axioms for linear discrete time.

3. $(K_i F K_i \varphi \wedge K_i F K_i (\varphi \rightarrow \psi)) \rightarrow K_i F K_i \psi$.

4. $K_i F^n K_i \varphi \rightarrow K_i F^{n'} K_i \varphi$, for all $n, n' \in \mathbb{N}$ such that $n < n'$.

5. $K_i F^n K_i \varphi \rightarrow K_i F K_i \varphi$, for all $n \in \mathbb{N}$.

6. $K_i F K_i \varphi \rightarrow \varphi$.

7. $K_i \varphi \rightarrow \neg F^n K_i \neg \varphi$, for all $n \in \mathbb{N}$.

8. $K_i \varphi \rightarrow F K_i K_i \varphi$.

9. $\neg K_i \varphi \rightarrow F K_i \neg K_i \varphi$.

Rules:

MP:
$$\frac{\varphi, \varphi \rightarrow \psi}{\psi}$$

RK:
$$\frac{\varphi}{K_i F K_i \varphi}, \quad \text{where } i \in AG.$$

Axiom 3 states, that if the agent has an algorithm to deduce[4] φ in some finite time and if he has an algorithm to deduce $\varphi \rightarrow \psi$ in some finite time, then this agent has also an algorithm to deduce ψ in some finite time.

Let us assume, that the agent requires n units of time to, when the need arises, deduce φ. We can then with complete certainty assume that the agent will be able to deduce φ when he will have more time at his disposal. This property is described by axiom 4, which states that if the agent can deduce φ in time not exceeding n units, then he can deduce φ in any time greater than n. Axiom 5 states that if the agent has an algorithm to infer φ in time not exceeding n units, then the agent has an algorithm to infer φ in a finite time. Axiom 6 is an equivalent to axiom **T**. Axiom 7 states that the agent's knowledge is consistent. Agents, in **LAK**, have the property of positive introspection, that is after deducing φ an agent may conduct an introspection of his knowledge and discover that he knows φ. Similarly with negative introspection. The properties of positive and negative introspections can be presented respectively with the use of axioms 7 and 8.

Let us observe that, in a proposed formalism, the implicit knowledge about formula φ is expressed with the use of $K_iFK_i\varphi$. For example, if φ is provable, then we can assume, that the agent has an algorithm to prove φ. On the basis of φ we can then deduce that $K_iFK_i\varphi$ (rule **RK**).

In **LAK** we can derive the rule:

Lemma 1.

Rule:

$$\frac{\varphi \to \psi}{K_iFK_i\varphi \to K_iFK_i\psi}$$

is a derived rule in **LAK**.

Proof. Let us suppose that $\varphi \to \psi$ is a theorem. By **RK** we can infer $K_iFK_i(\varphi \to \psi)$. The formula

$$K_iFK_i(\varphi \to \psi) \to (K_iFK_i\varphi \to K_iFK_i\psi)$$

is equvalent to postulate 3. So, the formula $K_iFK_i\varphi \to K_iFK_i\psi$ is inferred by **MP**. $\square$

In **LAK** we can prove the following statements:

1. $K_iFK_i\psi \to K_iFK_i(\varphi \to \psi)$
2. $(K_iFK_i\varphi \wedge K_iFK_i\psi) \leftrightarrow K_iFK_i(\varphi \wedge \psi)$
3. $K_iFK_i(\varphi \wedge \psi) \to K_iF(K_i\varphi \wedge K_i\psi)$
4. $K_iFK_i\varphi \to K_iFK_i(\varphi \vee \psi)$

Algorithmic knowledge can be applied to represent not only what the agent knows, or what he could know, but also to analyse a problem, how much time the agent needs to deduce – on the basis of his knowledge – that which he wants to know. The agent may not have knowledge about φ at a certain point in time. He might have however an inference procedure, which could lead him to obtain φ sentence in the future. The amount of time necessary to perform operations needed to infer a given sentence depends on a number of factors, out of which the most important seem to be: *complexity of the discussed sentence* and *the agent's computational power*. If the complexity of the discussed sentence and the agent's computational power are known, then the time needed by the agent to infer a given sentence may be estimated. In many cases not only the knowledge of the subject of what the agent knows is important, but the knowledge of the subject of what he does not know at a specified time is also important. Gaps in the agent's knowledge may limit his choices. His actions in such situations may be predictable, or appropriately explained.

Let us consider a rational agent who must perform some task in a specified time period. Let's also assume that finding a solution for completing his task is relatively easy, but finding an optimal plan is a very difficult problem, which is unsolvable in the allocated time period. In such a situation it is rational to find another, similar solution and give up searching for an optimal solution. A good example of such a situation is the problem of a traveling salesman. Calculating his route is quite an easy task, but selecting an optimal route is a very hard problem. Additionally, what is equally important, finding a solution which is close to an optimal solution might be a quite simple task. A rational agent then tries to select a route which is close to optimum and use this knowledge, instead of calculating an optimal route, calculating time of which might be very long.

Let us look at another example. When we use a public key to encrypt a message, we want to be sure that anybody who does not know the private key will be able to read the encrypted message. Although there are algorithms to obtain a private key on the basis of a public key, with the appropriate length of the key these algorithms work just long enough, that we might have, at least for now[5], a sense of maintained confidentiality in the case of the encrypted message. Our conviction that our encrypted message will not be quickly read is based on the complexity of the reasoning which should be performed in order to be able to read the encrypted message without the private key.

The lack of specific information can be deduced on the basis of other available information, with the use of certain assumptions about non-

contradiction of agents. For example, if the agent knows, that φ, then we can assume that the agent does not know that $\neg\varphi$, as long as the agent is reasoning in a non contradicting way. There is however another way, on the basis of which we can deduce the lack of algorithmic knowledge. The expectation that something cannot be known in specified time limits is based on the complexity of reasoning necessary to solve a given problem. Sending on encrypted message with the use of presently used cryptographic protocols, we can assume that our message will not be decrypted before it reaches an intended receiver, because the time necessary to deliver an electronic message is very short, while in order to break modern encoding protocols a few months are needed, and in some cases a few years. We can then rationally assume that the content of an encrypted message which we send will not be known by unauthorised persons for a few minutes.

When considering multi-agent systems we can ask a question: What can we say on the subject of the agents' meta-knowledge? What does agent j know about the knowledge of agent i? If we ask agent j a question: *How much time does agent i need to solve a question?*, then agent j will not have to use a test algorithm himself, does φ. In such a situation it will be enough for agent j to calculate the complexity of an algorithm needed to infer φ, in order to provide the information that agent i needs n time units. With the use of $\mathfrak{L}_{\mathbf{LAK}}$ language it is possible to model the knowledge of rational agents, who are not endowed with logical omniscience. Agents are rational because they are able to derive the logical consequences of what they know explicitly. On the other hand, because the agents have limited resources of knowledge, they cannot derive all the consequences of their knowledge. The explicit knowledge of the rational agents is not shut on any logical law, so the problem of logical omniscience in this case does not exist.

The deliberations which we have conducted on the subject of algorithmic knowledge have focused on a single moment of time. A formula of the type $K_i F^n K_i \varphi$ is a formula which talks about an agent's current capability to process knowledge. But this formula does not say anything about the agent's knowledge n time units from now (except in the case when $n = 0$). An interesting problem is the relation of the agent's algorithmic knowledge, at different moments in time, with respect to his algorithmic knowledge at other moments. Can we assume that the agent's explicit knowledge will hold in time? How do the agent's capabilities to reason change over time? If at the present moment in time we state that the following sentence is true: agent i needs n time units to calculate φ, will this sentence remain true in the future? Is the formula $K_i F^n K_i \varphi \rightarrow F K_i F^n K_i \varphi$ true? The answer to these questions is negative. Such a formula cannot be an axiom of a logical

system describing the knowledge of real agents, because we cannot assume its truthfulness for any formula φ. This is not a universal truth, because the logical value of φ may not be the same at different moments in time. What's more, in the future agent i may need more time to solve the same problem. In case when the agent under consideration is a man, accepting the assumption that cognitive or inference capabilities change over time seems to be justified. At certain moments in time an agent might be rested, more concentrated and need less time to decide φ at others, when he is tired, distracted, stressed-out, to settle the same question he might need more time. The same situation applies when we deal with artificial agents (processors, computers, robots, etc.). In the case of artificial agents we do not examine fatigue, stress or lack of concentration as factors causing changes in cognitive capabilities, but we look at for example an increase in the number of parallel running processes, reduced amount of available memory, etc.

In the case of classic logical systems which are used for formalisation of algorithmic knowledge, we accept usually the assumption that the computational capabilities of agents are not reduced over time. With such an assumption, the earlier considered formula might be true, if φ has a correct syntactic structure (for example φ is objective and does not contain a negation mark). We can ask the question, in what circumstances is such a formula true, or might some default rule of deduction be used for the evolution of the explicit knowledge in time? The systems mentioned earlier do not describe, however, the algorithmic knowledge of real agents, who have cognitive capabilities which change over time and have limited computational power and memory. These systems describe the knowledge of idealized agents who have no limit on resources and possess ownership of logical omniscience.

The main problem of single or multi-agent systems is however the description of real, implementable agents, namely ones that at the present stage of knowledge might be constructed or at least an idea of their construction (implementation) is known. This objective cannot be achieved if the appropriate limits of agents' knowledge are not properly included. Modal epistemic logics are not adapted to formalisation of reasoning regarding limits on knowledge resources, while the presented logic of algorithmic knowledge fulfils this objective.

N O T E S

[1] Speaking about logical omniscience here we have in mind the knowledge of all logical consequences resulting from the agent's explicit knowledge, as well as the knowledge of all logical laws.

[2] This quiz show aired on Polish Television some time ago.

[3] Algorithmic knowledge is connected then with the algorithm which creates it. It represents not only current knowledge, but also some sort of procedural knowledge.

[4] We do not mean just deducing in a logical sense. The algorithm may be a computational algorithm or some method of obtaining knowledge φ.

[5] When quantum computers will be accessible, modern cryptology, which is based on factorisation of large prime numbers, will cease to be applicable. The speed of operation of quantum computers, due to the simultaneous performance of a large number of computing operations, will be so fast that they will be able to perform factorisation of large prime numbers in a very short time. It is estimated that solving a problem, which for contemporary computers would take about 5000 years, for a quantum computer would take a dozen or so seconds.

R E F E R E N C E S

[1] K. Binmore, Shin H. S.,(1992,pp. 141-154), *Algorithmic knowledge and game theory*, Knowledge, belief, and strategic interaction (C. Bicchieri and M.-L. Dalla-Chiara, eds.), Cambridge University Press.

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

[3] J. Y. Halpern, Moses Y., Vardi M. Y., (1994, pp. 255-266) *Algorithmic knowledge, Theoretical Aspects of Reasoning about Knowledge*: Proc. Fifth Conference (R. Fagin, ed.), Morgan Kaufman.

[4] D. N. Ho, (1998, pp. 275-279), *On the epistemic foundations of agent theories.* In M. P. Singh, A. S. Rao, and M. J. Wooldridge, editors,Intelligent Agents IV. Proceedings of ATAL-97, volume 1365 of LNAI, Springer Verlag,

[5] D. N. Ho, *Knowledge, Logical Omniscience, and Resource-bounded Reasoning*, Institute for Information Systems, University of Luebeck, http://citeseerx.ist.psu.edu/viewdoc/downloaddoi=10.1.1.66.4838rep=rep1type=pdf

[6] A. Moreno, (2000), *Modelling rational inquiry in non-ideal agents.* Ph.d.thesis, Tarragona, Universitat Politecnica de Catalunya.

[7] D. Surowik, (2013), *Logic, Knowledge and Time* (in Polish), Bialystok University.