

# A Logic of Reasoning, Communication and Cooperation with Syntactic Knowledge

Thomas Ågotnes

Department of Informatics, University of Bergen  
PB. 7800, N-5020 Bergen, Norway

agotnes@ii.uib.no

Michal Walicki

Department of Informatics, University of Bergen  
PB. 7800, N-5020 Bergen, Norway

michal@ii.uib.no

## ABSTRACT

We present a general logic of explicit knowledge represented as finite sets of logical formulae which can evolve by non-deterministic reasoning and communication. It is partly based on Alternating-time Temporal Logic, which allows the expression of properties of cooperation. Properties of an agent's reasoning mechanism such as "the agent knows modus ponens" can be expressed. Instead of a common closure condition such as "if the agent knows both  $p$  and  $p \rightarrow q$ , he must also know  $q$ ", the following holds: "if the agent knows  $p$ ,  $p \rightarrow q$  and modus ponens, he has a strategy to get to know  $q$  in the future".

## Categories and Subject Descriptors

I.2.4 [Artificial Intelligence]: Knowledge Representation Formalisms and Methods—*Representations (procedural and rule-based)*, *Temporal logic*; I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence—*Intelligent agents, Multiagent systems*; I.2.3 [Artificial Intelligence]: Deduction and Theorem Proving—*Nonmon. reasoning and belief rev.*

## General Terms

Theory

## Keywords

Syntactic Knowledge Representation, Epistemic Logic, ATL

## 1. INTRODUCTION

We present a propositional logic of how the explicit knowledge of agents who represent their knowledge syntactically as formulae in a language can evolve over time. It is assumed that the storage is *finite*, i.e. that only finitely many formulae can be stored at the same time. We call the finite set of formulae known by an agent at a particular point in time the agent's *epistemic state*. Instead of assuming closure or consistency conditions on epistemic states, like closure under

logical consequence in modal epistemic logic, we assume that an agent at a point in time in addition to a finite epistemic state has a *mechanism* which can be used to transform the epistemic state into a new one. The agents are not necessarily *infallible* – the mechanisms may be unsound. In our logic, properties of epistemic states are expressed as knowledge of formulae while properties of mechanisms are expressed as knowledge of deduction rules. In addition to reasoning, agents in a multi-agent system can change their knowledge by *communication*. We are interested in agents who can make different *choices* about how to reason and communicate. Agents who can reason and communicate can cooperate strategically to e.g. reach certain epistemic states. Our logic is based on *Alternating-time Temporal Logic (ATL)* [4], which models such non-deterministic cooperation. ATL introduces cooperation modalities for sets  $G$  of agents: the formulae  $\langle\langle G \rangle\rangle \phi$ ,  $\langle\langle G \rangle\rangle \Box \phi$ ,  $\langle\langle G \rangle\rangle \mathcal{F} \phi$  and  $\langle\langle G \rangle\rangle \phi \mathcal{U} \psi$  means that the coalition  $G$  can cooperate – or that they have a strategy – to ensure that  $\phi$  is true in the next state, all future states, some future state and until  $\psi$  is true, respectively. The ATL language is interpreted in *concurrent game structures*.

## 2. LOGIC OF STATES AND MECHANISMS

An *epistemic state*  $s$  is a finite, but otherwise arbitrary, set of formulae in an *object language*  $OL$ , here assumed to be propositional logic:  $s \in \wp^{fin}(OL)$ . A *mechanism* models how the epistemic state of the agents can change over time. We assume that each agent at each time step sends a finite (possibly empty) set of formulae to each of the agents in the next time step including to himself. The latter is used to model reasoning. This action, by a particular agent, can be described by a tuple  $(s'_1, \dots, s'_n)$  with one finite set  $s'_j \in \wp^{fin}(OL)$  of object formulae for each agent  $j$ . We also assume that the agent can have several different *choices* of tuples to use. Since the agent cannot discern between situations in which he has the same epistemic state, the possible choices must be a function of the epistemic state. Thus, we model a mechanism for agent  $i$  as mapping an epistemic state  $s_i$  to a non-empty set  $R_i(s_i)$  of tuples. Note the different status of the members of  $(s'_1, \dots, s'_n) \in R_i(s_i)$ :  $s'_i$  is the set  $i$  wants himself to know (maybe just remember from last time step, if  $s'_i \subseteq s_i$ ) in the next time step, while  $s'_j$  ( $j \neq i$ ) is the information  $i$  sends to another agent  $j$ . A *global state* in the multi agent system consists of one epistemic state for each agent. The transition between global states is defined as follows: each agent  $i$  with ep. state  $s_i$  selects a tuple  $(s'_1, \dots, s'_n) \in R_i(s_i)$ , and an agent's epistemic state in the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

AAMAS'05, July 25-29, 2005, Utrecht, Netherlands.  
Copyright 2005 ACM 1-59593-094-9/05/0007 ...\$5.00.

$$\left( \Delta_1 \{p\} \wedge \Delta_2 \{p \rightarrow q\} \wedge \langle \langle \emptyset \rangle \rangle \square \left( \tilde{\Delta}_{12} \left\{ \frac{t \sqcup \{a\}}{\{a\}} \right\} \wedge \tilde{\Delta}_{22} \left\{ \frac{t \sqcup \{a, a \rightarrow b\}}{t \sqcup \{b\}} \right\} \wedge \tilde{\nabla}_{22} \left\{ \frac{t}{t \sqcup u} \right\} \right) \right) \rightarrow \langle \langle \{1, 2\} \rangle \rangle \mathcal{F} \Delta_2 \{q\}$$

**Figure 1: A valid formula. The ATL formula  $\langle \langle \emptyset \rangle \rangle \square \phi$  says that  $\phi$  holds in all future states.**

next global state is the union of the sets sent to that agent from all agents (including from himself) – the next global state is  $(\bigcup_{i=1}^n s_1^i, \dots, \bigcup_{i=1}^n s_n^i)$ . This definition of communication, where an agent can “insert” formulae directly into another agent’s epistemic state, are useful in systems with reliable communication – but also in other applications since the language presented shortly can express *restrictions* on the *form* of the formulae that can be inserted. The satisfiability relation for the logical language is defined with respect to a mechanism for each agent and a global state. Technically, this is done by translating a mechanism to a concurrent game structure.

The logical language is the language of propositional logic extended with three new fragments. First, *epistemic operators* are used in statements about the agents’ current knowledge. The fact that agent  $i$  *knows at least*, *knows at most*, and *knows exactly* a finite set  $X$  of *OL* formulae is expressed by the formulae  $\Delta_i X$  ( $i$ ’s epistemic state includes  $X$ ),  $\nabla_i X$  ( $i$ ’s epistemic state is included in  $X$ ), and  $\diamond_i X \equiv \Delta_i X \wedge \nabla_i X$ , respectively. The epistemic operators are further discussed in [2]. Second, ATL operators are used to express statements about strategic cooperation. Third, *rule operators* and *rules* are used in statements about the agents’ abilities to reason and communicate, i.e. about their mechanisms. Rules are defined using two types of formal variables:  $a, b, \dots$  are used as place-holders for formulae, and  $t, u, \dots$  for finite *sets* of formulae. A rule consists of an *antecedent* and a *consequent*; both representing finite sets of formulae — possibly containing variables and set building operators  $\sqcup$  and  $\sqcap$ . Examples of rules are:

$$R_1 = \frac{t \sqcup \{a \rightarrow b, a\}}{t \sqcup \{b\}} \quad R_2 = \frac{t}{t \sqcup u} \quad R_3 = \frac{t \sqcup \{a\}}{\{a\}}$$

A rule is interpreted as a relation between sets of formulae simultaneously matching the antecedent and consequent. Formulae  $\tilde{\Delta}_{ij} R$  and  $\tilde{\nabla}_{ij} R$  are used to express that agent  $i$  knows at most and at least, respectively, the set of rules  $R$  for communication to agent  $j$ . Semantically, if the current epistemic state of agent  $i$  is  $s$ , both formulae are statements about the set  $S_j^i(s) = \{s_j^i : (s_1^i, \dots, s_n^i) \in R_i(s)\}$  of sets agent  $i$  can send to agent  $j$ .  $\tilde{\Delta}_{ij} R$  says that when  $s$  matches the antecedent of a rule in  $R$  then  $S_j^i(s)$  includes the consequent;  $\tilde{\nabla}_{ij} R$  that every set in  $S_j^i(s)$  must be the consequent in some rule in  $R$  with an antecedent matching  $s$ . While the same syntax is used for reasoning ( $i = j$ ) and communication ( $i \neq j$ ), the actual rules used in the two cases will typically be different. For example,  $\tilde{\Delta}_{ii} \{R_1\}$  expresses that  $i$  can reason with modus ponens (MP),  $\tilde{\nabla}_{ii} \{R_2\}$  that  $i$  must reason monotonically (since the consequent extends the antecedent), while  $\tilde{\Delta}_{ij} \{R_3\}$  and  $\tilde{\nabla}_{ij} \{R_3\}$  express that  $i$  can communicate any formula  $i$  knows to  $j$ , and that everything  $i$  can communicate to  $j$  must be known by  $i$ , respectively. We also let  $\tilde{\diamond}_{ij} R \equiv \tilde{\Delta}_{ij} R \wedge \tilde{\nabla}_{ij} R$ .

An example involving several of the operators is shown in Fig. 1. The formula says that if agent 2 knows MP and the formula  $p \rightarrow q$  and reasons monotonically, and agent 1 knows the formula  $p$  and can communicate any known formula to 2, then agents 1 and 2 can cooperate to make 2 know the formula  $q$  in the future. Some general syntactic properties of the logic are:

1.  $\tilde{\diamond}_i X \rightarrow \left( \tilde{\diamond}_{ij} R \leftrightarrow \langle \langle \emptyset \rangle \rangle \square (\tilde{\diamond}_i X \rightarrow \tilde{\diamond}_{ij} R) \right)$
2. When  $i \notin \Gamma$ :  $\langle \langle \Gamma \rangle \rangle \circ \Delta_j X \rightarrow \neg \langle \langle \{i\} \rangle \rangle \circ \neg \Delta_j X$
3. When  $\Gamma \cap \Gamma' = \emptyset$ :  $(\langle \langle \Gamma \rangle \rangle \circ \Delta_j X \wedge \langle \langle \Gamma' \rangle \rangle \circ \Delta_j Y) \rightarrow \langle \langle \Gamma \cup \Gamma' \rangle \rangle \circ \Delta_j (X \cup Y)$

### 3. CONCLUSIONS

We have presented a general logical framework for modelling how the explicit knowledge of agents who represent knowledge syntactically in a logical language can evolve as a result of reasoning and communication. We argue that the proposed language is expressive enough to capture theoretically interesting properties, and that the unified model of reasoning and communication is applicable in many circumstances. Further details can be found in [1].

*Konolige’s deduction model* [6], *active logics* [5] and *timed reasoning logics* [3] also model agents with sets of known formulae and transition relations described by rules. These approaches, however, describe a linear and deterministic, instead of a branching and non-deterministic, future, and have no concept of strategic cooperation.

Possible future work include a complete axiomatization.

### 4. REFERENCES

- [1] T. Ågotnes. *A Logic of Finite Syntactic Epistemic States*. PhD thesis, Department of Informatics, University of Bergen, 2004.
- [2] T. Ågotnes and M. Walicki. A logic for reasoning about agents with finite explicit knowledge. In B. Tessem, P. Ala-Siuru, P. Doherty, and B. Mayoh, editors, *Proc. of the 8th Scandinavian Conference on Artificial Intelligence*, Frontiers in Artificial Intelligence and Applications, pages 163–174. IOS Press, Nov 2003.
- [3] N. Alechina, B. Logan, and M. Whitsey. A complete and decidable logic for resource-bounded agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2004)*, pages 606–613. ACM Press, Jul 2004.
- [4] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time temporal logic. In *38th Annual Symposium on Foundations of Computer Science*, pages 100–109, Miami Beach, Florida, 20–22 Oct. 1997. IEEE.
- [5] J. Elgot-Drapkin, S. Kraus, M. Miller, M. Nirkhe, and D. Perlis. Active logics: A unified formal approach to episodic reasoning. Techn. Rep. CS-TR-4072, 1999.
- [6] K. Konolige. *A Deduction Model of Belief*. Morgan Kaufmann Publishers, Los Altos, California, 1986.