Electronic Notes in Theoretical Computer Science 70 No. 5 (2002)
URL: http://www.elsevier.nl/locate/entcs/volume70.html 5 pages

Logics and Multi-agents: towards a new symbolic model of cognition

Paolo Torroni [1,2]

Dipartimento di Elettronica, Informatica e Sistemistica
Università di Bologna
V.le Risorgimento 2, 40136 Bologna, Italy

Extended Abstract

The last edition of CLIMA, held in 2001 in Paphos (Cyprus) ended with a panel session on the role of Computational Logic (CL) in Multi-Agent Systems (MAS). Two dimensions in MAS development were singled out and discussed: on the one hand reactivity vs. rationality, and on the other hand individuals vs. societies. Most of the points discussed aimed at justifying and motivating the application of CL techniques to MAS development: should be logics used to implement the individuals, or the society, or both? should be logics used to model the reactive part, or the pro-active part, or both? what do we want to achieve in terms of properties, openness to integration, etc.?

A most intuitive reply to these questions could be that logic should be used for what logic is good at. For instance, logic programming-based techniques such as abductive and inductive logic programming seem suitable for modelling agent hypothetical reasoning and adaptability. Modal logic operators such as those adopted by a BDI agent model [3] could be a powerful and synthetic way to describe the agent behaviour and to put it into relationship with the other agents in a society. Model checking-based techniques can be applied to the verification of agent systems. A combination of multiple approaches, like modal and temporal logics, or abduction and induction in a logic programming framework, could be the key to achieve a more comprehensive agent and agent system architecture. But in this case, to determine which properties of the chosen combinations hold is not an easy task.

At the time of this new edition of CLIMA, while the debate about the role of CL in MAS is still open, from within the CL community we are witnessing a growth of interest for Multi-Agent Systems considered *per se* as an interesting

[1] I would like to thank Evelina Lamma, Francesca Toni, and Ken Satoh for fruitful discussions and comments about this document. This work has been partially supported by the SOCS project, funded by the CEC, contract IST-2001-32530.
[2] Email: ptorroni@deis.unibo.it

cognitive model. This is due to many reasons, among which, we would say, the need to put "abstract" reasoning in the context of a "concrete" environment, and to use logic not only to solve problems in a virtual world, but in a real arena. The multi-agent metaphor of intelligent individuals that are situated into dynamic and unpredictable environments and that can interact with each other by updating their beliefs, can be regarded then as the basis for a new symbolic model of cognition.

Logic-based agents and the outer world

Some recent work on Logic Programming outlines this new concept of intelligent system. In [1], Kowalski says: "it is the objective perspective of multiagent systems that forces me to acknowledge the existence of a real environment, which exists independently of individual agents: As I see it now, if there is only one agent, then that agent's environment might only be virtual. But if there are several agents interacting with one another, and if all of them are equally real, then the environment of each agent must include the other agents, and therefore that environment itself must also be real. This real environment, shared by several agents, can be understood as a classical model-theoretic, semantic structure. It gives meaning to the agents' thoughts, making some thoughts true and other thoughts false. It grounds their thoughts in reality."

It is our opinion that much of the work presented in this workshop well reflects this concept. Speculative computation on the one hand and planning together with action execution on the other hand reconcile the agent reasoning with the effect of actions made on an external world. Techniques proposed to deal with message loss or modification clearly picture the idea of an environment where logic based agents are situated that could indeed be very different from the model that they have of it. The introduction of hierarchies, roles, and preferences puts the agent in a context, which sometimes we can call society, or institution, which is at a higher level than that of the single individuals, and which gives a meaning to the agents' thoughts and behaviour.

Why is this new model interesting from a CL perspective? Taking into account an environment with its own semantics means to accept destructive assignment [1]. If we consider MAS as a distributed and concurrent computational system, considering multiple autonomous agents could imply imposing a committed choice at every step. If we want to adopt this new paradigm for Computational Logic, what are the choices that we ought to do? What new assumptions should we make, and on the other hand, how could we accommodate these new features in our background?

If we consider an agent's viewpoint, the other agents in the system could be seen as a part of the environment. Therefore, an aspect that is fundamental in this new cognitive model is that of communication, since it is one of the ways agents become aware of each other. It is inter-agent communication that could make of a system of agents with symbolic knowledge representation a system for collaborative problem solving, for instance by influencing other agents' mental states [2]. But accommodating communication in a logic-based model of agent is not a trivial task, and has several semantic implications. For instance, if we use a form of abduction to model communication and model a communicative act as an abducible predicate [4,5], what is then the semantics of such predicate in a multi-agent context? Is it still an abducible predicate, or does it become a fact once it is transmitted?

In general, the question could be put in the following way: how to accommodate in an agent knowledge representation and reasoning activity the external inputs given by a dynamically evolving environment?

Reconciling individual consistency with collective consistency

The problem raised above is tightly related to another central issue, also considered by some papers in this workshop, that of consistency. Putting together several agents with different knowledge bases could indeed lead to system inconsistency, depending for instance on what semantics we want to give to the overall system, and on the presence of integrity constraints. Constraints can be seen at an abstract level as formulas that must be true at all times, and they can be used in practice to express the agent's behaviour and interaction protocols.

Let us mention a couple of concrete application scenarios. In distributed systems management there might be agents that provide services (resource managers), and they may have some policies about the access to such resources. Client agents that request services to the resource managers may also have their own policies that constrain and rule their requests. Integrity constraints on the agent's behaviour can be used to describe such policies. In a similar way it could be possible to model, for instance, electronic institutions, where a "governor" agent could encode the institution's norms in the form of integrity constraints.

We have shown two examples of systems where integrity constraints play a key role, both at the individual's level and at the system's level. In the first case, they implicitly define interaction protocols, as resulting from the composition of individuals; in the second case they explicitly encode behaviour rules and norms with which everyone in an institution must comply. When we put together different individuals it becomes important to define what is consistent and what is not. In the abstract, the question becomes: what semantics could we give to a system of logic-based agents? What does it mean to preserve the individual's and the system's integrity? How to maintain the consistency of the overall system emerging from the "composition", or better, interaction, of multiple and independent interpretations of the world?

Computationally practical BDI models

There is a last issue that we would like to address. It is often the case that new proposals are made with the ultimate purpose to have a system that is implementable, and which at the same time has properties which is possible to determine and to prove (this is one of the claimed advantages of a logic-based approach). But in fact, what we witness is a real problem in bridging the gap between theory and practice. Which agent architectures could we adopt to the purpose?

Indeed, a reference architecture for the agent behaviour is the well known BDI model, based on Beliefs, Desires, and Intentions, and its variations and evolutions. After more than a decade from its introduction, there is still a considerable amount of work being done and to be done, aimed at giving links between computational logics and architectures for BDI, at verifying whether practical implementations of BDI do actually meet the theoretical requirements or whether a BDI agent adapts itself to a particular BDI-strategy, at recasting the foundations of BDI into a logic programming framework, at providing proof methods to establish the consistency of classes of formulas to represent introspective beliefs. BDI seems then a powerful way to model the agent behaviour and the evolution of a society of agents. The idea of possible worlds is indeed appealing as a possible representation of an evolving environment. But to what extent are applications of full BDI proof-theoretic? Then, the question is: will we get to a comprehensive implementation of a BDI agent or is this a utopia? What simplifications to this model can be considered acceptable in a realistic application?

Conclusion

Much of this work has been inspired by the papers presented at the workshop. The aim of this contribution was to raise some questions which we consider central in the area of logic-based multi-agent systems, and to leave them as open topics of discussion. The questions are about the relationship between the internal reasoning of agents and the outer world, about the semantics of a system composed of agents that reason on a private knowledge, and about the implementation of a BDI architecture. It is our belief that answers to them could promote a significant advance in both the Multi-Agent Systems and Computational Logic research of the next years.

References

[1] Kowalski, R. A., *Logic programming and the real world*, Logic Programming Newsletter (2001).

[2] Panzarasa, P., N. R. Jennings and T. J. Norman, *Formalising collaborative decision making and practical reasoning in multi-agent systems*, Journal of Logic

and Computation **12** (2002).

[3] Rao, A. and M. Georgeff, *An abstract architecture for rational agents*, in: *Proceedings of the International Workshop on Knowledge Representation (KR'92)*, 1992.

[4] Sadri, F., F. Toni and P. Torroni, *Dialogues for negotiation: agent varieties and dialogue sequences*, in: *Intelligent Agents VIII, ATAL'01 Revised Papers*, LNAI **2333** (2002), pp. 405–421.

[5] Satoh, K. and K. Yamamoto, *Speculative computation with multi-agent belief revision*, in: *Proceedings AAMAS, Bologna, Italy* (2002), pp. 897–904.