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

Paolo Torroni

DEIS, Università di Bologna
Viale Risorgimento 2, 40136 Bologna, Italy
ptorroni@deis.unibo.it

Abstract. In the last years, Computational Logic proved to be a successful approach to several aspects of Multi-Agent Systems design. Some examples of it are logic programming-based agent reasoning and model checking-based verification techniques, applied to agents and agent systems. At the same time, from the Computational Logic side we are witnessing a growth in the interest for Multi-Agent Systems. Some recent directions of research seem to push towards a new idea of intelligent systems, and the metaphor of intelligent individuals that are situated into dynamic environments and that can interact with each other, updating their mutual beliefs, is being regarded as the basis for a new symbolic model of cognition. It is our intention to propose some open questions about this new perspective to warm up a discussion panel for CLIMA'02. It is our belief that answers to them could foster a significant advance in both the Multi-Agent Systems and Computational Logic research of the next years.

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 behavior 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 pave the way to a more comprehensive agent and agent system architecture, while bringing about on the other hand more issues such as how to determine which properties of the chosen combinations do hold.

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 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.

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 agents in a context that gives itself meaning to the agents' thoughts.

It would be interesting to investigate and to exchange opinions about the reasons why this new model is interesting. 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 a system of agents with symbolic knowledge representation be 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?

This 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. What semantics could we give to a system of logic-based agents? How to maintain the consistency of the overall system emerging from the "composition", or better, interaction, of multiple and independent interpretations of the world?

A last issue that we would like to put forward is which agent architectures could we adopt to our purposes. Indeed, a reference architecture and model 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 much work, 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. The BDI seems then a powerful way to model the agent behavior 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?

We would like to conclude this by resuming the questions raised above, and leaving them as open topics of discussion:

– How to accommodate communication in a logic based model of agent? How to bind the product of a logic derivation with the effects that it has on the external world? How is it possible to do it within logics?
– What does it mean to join several individuals that reason on a private knowledge? What does it mean to preserve the individual's and the system's integrity? What can it be the role of logic and logic-based techniques?
– What do you see as the future for agent models? Will we get to a comprehensive implementation of a BDI agent or is this a utopia? What simplifications to the model can be considered acceptable in a realistic application?

References

1. R. A. Kowalski. Logic programming and the real world. *Logic Programming Newsletter*, 2001.
2. P. Panzarasa, N. R. Jennings, and T. J. Norman. Formalising collaborative decision making and practical reasoning in multi-agent systems. *Journal of Logic and Computation*, 12(1), 2002.
3. A. Rao and M. Georgeff. An abstract architecture for rational agents. In *Proceedings of the International Workshop on Knowledge Representation (KR'92)*, 1992.
4. F. Sadri, F. Toni, and P. Torroni. Dialogues for negotiation: agent varieties and dialogue sequences. In *Intelligent Agents VIII, 8th International Workshop, ATAL 2001 Seattle, WA, USA, August 1-3, 2001 Revised Papers*, volume 2333 of *LNAI*. Springer Verlag, 2002.
5. K. Satoh and K. Yamamoto. Speculative computation with multi-agent belief revision. In *Proceedings AAMAS, Bologna, Italy*, 2002. Accepted for publication.