An architecture for rational agents interacting with complex environments

A. Stankevicius M. Capobianco C. I. Chesñevar

Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
e-mail: {ags,mc,cic}@cs.uns.edu.ar

1 Introduction

Artificial Intelligence encompasses the enterprise of building an *intelligent agent.* To this purpose, the notions of *agency* and *intelligence* need to be precisely defined. Much to our demise, intelligence has proved to be hard to grasp in a formal way. To overcome this, Stuart Russell—among others—propounded *rational agency* as an alternative characterization for intelligent agency. In his own words [9], "rational agents [...] are agents whose actions make sense from the point of view of the information possessed by the agents and its goals". Simply put, a rational agent is an agent that does the right thing based on its beliefs [5]. Naturally, the construction of rational agents is by no means an easy assignment; in fact, it is a large and complex endeavor.

Software Engineering has a long standing tradition tackling tough projects. As a general norm, proper modularization has provided the key to success in these tasks. Therefore, different *agent architectures* have been proposed. The agent architecture supplies a separation of concerns, inducing a tentative modularization. However, at the present state of affairs, to build a fully rational agent is rather unattainable; several issues remain overlooked, or just partially addressed (*i.e.*, addressed under too unrealistic assumptions). Rao and Wooldridge [7] underscore the following aspects:

> *Rational agents are not solitary entities living in an uninhabited static environment; they are embedded in a continuously changing environment and have to constantly interact with other agents. Hence, communication with other agents and interaction with the environment are key concerns within this field....*

In this paper we sketch an agent architecture suitable to be used as a tool for exploring agent perception and multiagent interaction. Nowadays, there is a tenuous link between the theoretical work in rational agents and their implementation. It is our intention to reach a good trade-off between expressiveness and implementability.

Argumentative frameworks [2] constitute an appropriate formalism for knowledge representation and reasoning (mainly) as a consequence of its ability to deal with incomplete and/or contradictory information. Moreover, most of these systems have been successfully implemented without neglecting their expressiveness. For this reason, an argumentative system underpins our architecture in order to represent the agent's knowledge, and conclude from it. The remainder of this paper is structured as follows: section 2 delineates the proposed architecture, and discusses its components. Finally section 3 highlights the related work.

2 Agent architecture

Although consensus cannot be easily reached, we believe that a rational agent should be able to pass the following criteria:

- Reactivity: the ability to sense and act selectively.
- Autonomy: goal-directed, proactive, and self-starting behavior.
- Inferential capabilities: the agent reasons from its knowledge and goes beyond the information given; it may have explicit models of itself, the environment, and/or other agents.
- Social capabilities: the ability to cooperate and/or coordinate with other agents to achieve common goals, bargaining with them if necessary.

Any architecture for rational agents is likely to accommodate many of these features. The architecture we envision (portrayed in Figure 1) is grounded on these four pillars. In the sequel, we detail the role of each component.

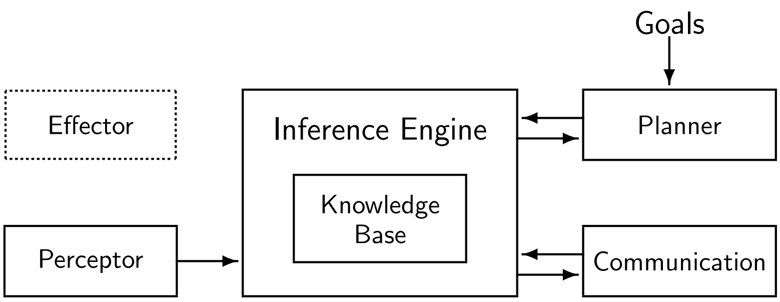

Figure 1: An sketch of the agent architecture.

2.1 Knowledge base

The knowledge base is the information repository of the agent. Even though multi-modal logics are usually the first choice for modelling an agent's mental state, the expressiveness of these approaches tend to be lost in the transition towards practical systems [7]. On the contrary, adopting a logic program as a model for the epistemic state of an agent admits switching back and forth between theory and practice seamlessly.

Unfortunately, conventional logic programming cannot deal with partial and potentially contradictory information [3], a recurring situation when modelling practical agents. This objection endorsed the development of extensions to logic programming that deal with these inconveniences. Observation based defeasible logic programming (ODeLP) [1] is a formalism recently developed to model the epistemic state of agents in dynamic domains. ODeLP combines the advantages of logic programming with defeasible argumentation and perception mechanisms. Accordingly, we have chosen the language of ODeLP to conform the knowledge base of our agent.

2.2 Inference engine

To draw a conclusion from an ODeLP program arguments for and against a certain thesis are formed. These arguments undergo a dialectical analysis that establishes their current statuses. Hence, a query q can obtain as answer YES, if there exists a justified argument for q, NO, if there exists a justified argument for the complement of q, or UNDECIDED, if neither q nor its complement are supported by a justified argument. In our architecture, the inference engine accesses ODeLP program functioning as the knowledge base to solve queries posed by the other units.

2.3 Planning unit

The planning unit exercises the control of the agent. It constructs plans to achieve agent's goals based upon the current epistemic state. In this setting, we understand a plan to mean any sequence of actions. The basic actions available to the agent involve consulting other agents about their epistemic state, engaging them in negotiations, and any other elementary operation performed by its effectors.

2.4 Communication unit

The communication unit embodies the *high-level* perceptors and effectors that manage the interaction among agents. Two levels can be identified. The first one, *atomic interaction*, includes both querying agents to gain insights about their epistemic states, and accomplishing queries from other agents. In contrast *composite interaction* comprehends any non-atomic interaction among agents, such as coordination, cooperation, and negotiation.

Atomic interaction is easily achieved: queries can be posed by means of the communication unit, and the inquiries received can be easily answered through the inference engine. We assume that every agent is aware of the existence of its counterparts: the communication unit maintains the information required to access any of them. Note that atomic interaction among agents exchanges unstructured information only. Thus, no specific protocol is required.

For composite interaction, we focus our attention on deliberation. We endorse conceiving most of the other high level interactions (such as cooperation, persuasion, negotiation, etc.) as simple by-products of a carefully conducted deliberation. To make it precise, by deliberation be understand the process through which a set of entities come to mutually accepted position regarding a given issue. For instance, deliberation can in turn be seen as a persuasion conduit, where the agent prevailing the dispute influences the epistemic state of the other party. In contrast to atomic interaction, the deliberation process involves the exchange of elaborated evidence about the topic being considered. For instance, in this context the parties exchange arguments either supporting or rebutting a given claim. Clearly, a protocol governing these transactions need be defined.

2.5 Perceptors and effectors

A rational agent residing in a complex dynamic environment must be able to gather information by sensing its surroundings [5]. The perceptor unit is responsible for handling this interaction, reporting environment changes so they can be incorporated into the knowledge base. The effector unit executes a set of actions over the world, as to follow the requirements of the planner. This unit often owns a variety of actions. As we concentrate on agent perception and multi-agent negotiation, the architecture does not require any particular action besides those provided by the communication unit. Consequently, Figure 1 depicts the effector unit in dashed lines.

3 Related work

In this section we survey the main ideas on agent perception and multi-agent negotiation being developed using the proposed architecture. These results comprise the research being conducted in two different, but interconnected, research lines. In what follows, we explore each topic individually.

3.1 Agent perception

We already stressed the importance of considering a rational agent as an entity interacting in a changing environment. Clearly, it is impossible to provide the agent with all the required information from its conception; this prompts for the definition of mechanisms suitable for agent perception. Our perceptor unit senses the environment, providing concrete facts about the world. However, the agent cannot monitor the entire state of the world at all time; the best perception can do is to provide the agent with a sort of sampling at discrete times, or over shot time intervals [5].

Despite of this, there is a subtlety that complicates perception: the world, as it evolves, need not be consistent with its previous states (*i.e.*, a new perception may contradict a previous one). An agent sensing its environment must be able to deal with the inconsistencies arising from its perceptions.

The argumentative framework encoding the agent's knowledge must be able to distinguish its previous knowledge from the new information acquired through perception. Consequently, we have provided the system of ODeLP with the ability to update the knowledge base as it receives new perceptions. The proposal developed in [1] addresses all the aforementioned issues.

3.2 Interaction among agents

The traditional approach for modelling multiagent interaction resorts to game theory [8]. Several insightful issues have been explored under this conception. Unfortunately, it usually underpins on a strong assumption: each agent should be aware of the complete pay-off matrix *before* the interaction even starts. The implications of this strong assumption undermines the applicability of the whole approach.[1]

In a prominent paper, Parsons *et al.* [4] develop the idea that multiagent interaction can be seen as an argumentative process. In an attempt to follow their intuitions, we also adopt

[1]it should be noted that there also exists variants where the matrix can be dynamically discovered.

an argumentative framework to model multi-agent negotiation (note that an argumentative framework is already present in our architecture).

Recent progress in defeasible argumentation favors dialectical characterization. For instance, Prakken's dialogue game [6], Simari's dialectical trees [10], etc. Under this particular view, two parties take opposing positions regarding a certain matter, and alternatively pose reasons backing their stances. The likeness with a set of agents interacting with each other is striking.

We convey that, instead of circumscribing solely to the actual interaction, the entire process should be taken into account. At the current state of the research, the following steps have been identified:

1. An agent decides that it needs to interact with another agent. The planning unit is responsible for taking this decision.

2. The agent engages the chosen counterpart. Once the communication unit contacts the other agent, the interaction is ready to begin.

3. The actual interaction takes place. In this step, the interaction is conducted according to an argumentative protocol.

4. The outcome of the interaction is accounted. Note that when the agent "loosing" the negotiation accepts the matter being debated, its knowledge base becomes contradictory. In this step, the ability to deal with contradictory information becomes indispensable.

It goes without saying that these steps require further refinement. In [11] steps 3 and 4 are throughfuly discused.

References

[1] Capobianco, M., Chesñevar, C. I., and Simari, G. R. An argumentative formalism for implementing rational agents. In *Proceedings del 7mo Congreso Argentino de Ciencias de la Computación (CACIC)* (2001), Universidad Nacional de la Patagonia Austral, pp. 1051–1062.

[2] Chesñevar, C. I., Maguitman, A., and Loui, R. Logical Models of Argument. *ACM Computing Surveys 32*, 4 (2000), 337–383.

[3] Gelfond, M., and Lifschitz, V. Logic Programs with Classical Negation. In *Proceedings of the 7th International Conference on Logic Programming* (June 1990), D. H. D. Warren and P. Szeredi, Eds., pp. 579–597.

[4] Parsons, S., Sierra, C., and Jennings, N. Agents that reason and negotiate by arguing. *Journal of Logic and Computation 8*, 3 (1998), 261–292.

[5] Pollock, J. L. Rational Cognition in OSCAR. In *Proceedings of the 6th International Workshop on Agent Theories, Architectures, and Languages* (1999). To appear.

[6] Prakken, H. *Logical Tools for Modelling Legal Argument*. Kluwer Academic Publishers, 1997.

[7] Rao, A., and Wooldridge, M. Foundations of Rational Agency. In *Foundations of Rational Agency*, no. 14 in Applied Logic Series. Kluwer Academic Publishers, 1999, pp. 1–10.

[8] Rosenschein, J. R., and Zlotkin, G. *Rules of Encounter*. MIT Press, 1994.

[9] Russell, S. J. Rationality and Intelligence. *Artificial Intelligence 94*, 1–2 (1997), 57–77.

[10] Simari, G. R., Chesñevar, C. I., and García, A. J. The Role of Dialectics in Defeasible Argumentation. In *Proceedings of the XIV Conferencia Internacional de la Sociedad Chilena para Ciencias de la Computación* (Nov. 1994), pp. 111–121.

[11] Stankevicius, A. G., and Simari, G. R. An abstract model for the process of deliberation within multiagent systems. In *Proceedings del 7mo Congreso Argentino de Ciencias de la Computación (CACIC)* (2001), Universidad Nacional de la Patagonia Austral, pp. 1017–1026.