Rational Decision Making in Autonomous Agents

Gerardo I. Simari† [1] Simon D. Parsons‡

† Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA) [2]
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur. Av. Alem 1253, (8000) Bahía Blanca, Argentina
Tel: ++54 291 4595135 - Fax: ++54 291 4595136
gis@cs.uns.edu.ar

‡Department of Computer and Information Science
Brooklyn College, City University of New York
2900 Bedford Avenue, Brooklyn, NY 11210, USA
Tel: ++1 718 951 5520 - Fax: ++1 718 951 4842
parsons@sci.brooklyn.cuny.edu

Abstract

Making rational decisions is one of the key elements in the design of autonomous agents with successful behavior. Even though there have been many proposals for the support of decision making, most of them can be described either as *descriptive* or *prescriptive*. The main goal of our work is to establish the relationship between two of these models, namely BDI and MDPs, in order to gain further understanding of how decisions in one model are viewed from the point of view of the other. This goal is important for the development of agent design strategies that unite the best of both worlds.

1 Introduction and Background

The key to implementing successful behavior in autonomous agents is *deciding what to do next*; this is true for softbots playing computer games [1], or robots playing soccer [7, 2]. This problem has been widely studied, and a number of models that carry the name of *architectures* have been formulated. Most of these approaches fall into the following categorization:

- *Descriptive* approaches, which are based on analyzing the way that people or animals make decisions. These approaches include, for instance, the belief/desire/intention (BDI) approach [5] and the behavior-based approach [3]. For example, the PRS [8] model can be considered an architecture for BDI decision making, and the subsumption architecture [6] as an architecture for behavior-based decision making.

[1]Partially supported by *Comisión de Investigaciones Científicas, Gobierno de la Provincia de Buenos Aires*, and by the *Agencia Nacional de Promoción Científica y Tecnológica* (PICT 13096)

[2]Member of the IICyTI (Instituto de Investigación Científica en Ciencia y Tecnología Informática).

- *Prescriptive* approaches, which attempt to identify the *optimal* decision. They are typically based on decision theory [12], and one family of approaches within this class, which is currently the subject of much research interest, is the family of *Markov decision processes* (MDPs) [11].

Since BDI, and implementations thereof, have been widely used by agent developers, it is interesting to ask about the quality of the decisions that the model makes. It seems natural that this will depend upon the exact nature of the task, and this was experimentally validated by Kinny and Georgeff [9]. In particular Kinny and Georgeff showed that the performance of an agent depended upon the speed with which its environment changed, the amount of information the agent has at its disposal, and the likelihood of its actions having their intended effect.

Another of Kinny and Georgeff's findings was that the performance of the agent depended upon how often, broadly speaking, it considered whether it had made the right decision (its *commitment strategy* in the language of the BDI model). Following up on this, Schut and Wooldridge [13, 14, 16] considered a range of models for making this meta-level decision about whether the last decision was still a good one, even using an MDP model [15] to optimize it.

All of this work, however, has only been able to compare different commitment strategies with one another using a metric of how well the agent performs on the task rather than with any notion of what the optimum performance is. All that we know is that, as a heuristic approach, the BDI model is likely to be sub-optimal. We just don't know *how* sub-optimal. The trade-off, the reason we may be prepared to accept this sub-optimality, is that the BDI model is much more tractable than prescriptive approaches like MDPs. As we have shown [17], MDPs can be intractable even for rather small problems.

2 Heuristics vs. Decision-Theoretic Optimality

Our work builds on that of [15], which focuses on understanding the relationship between the BDI model and MDPs. One way we are investigating this is by examining how good a solution the BDI model produces in comparison with MDPs on the same testbed used first by Kinny and Georgeff and then by Schut and Wooldridge. It turns out that to apply MDPs on the testbed, we have to resort to some novel approximations [17]. In this section, we will briefly describe the TileWorld domain, and some of the results we have obtained from the comparison of models.

2.1 The TILEWORLD Domain

The TILEWORLD testbed [10] is a grid environment occupied by agents, tiles, holes, and obstacles. The agent's objective is to score as many points as possible by filling up holes, which can be done by pushing the tiles into them. The agent can move in any direction (even diagonally); the only restriction is that the obstacles must be avoided. This environment is *dynamic*, so holes may appear and disappear randomly in accordance to a series of world parameters, which can be varied by the experimenter.

Because this environment, though simple to describe, is too complex for most experiments, we adopted the simplified testbed used in [9, 13]. The simplifications to the model are: tiles are omitted, so an agent can score points simply by moving to a hole; agents have perfect, zero-cost knowledge of the state of the world; and agents build correct and complete plans for visiting a single hole (they do not plan tours for visiting more than one hole). This domain, although simplistic, is useful in the evaluation of the effectiveness of situated agents. One of its main advantages is that it can be easily scaled up to provide difficult and unsolvable problems.

2.2 A Comparison of Models

In an MDP, the world can be modelled by taking into account *every* possible action in *every* possible state. For the simplified TILEWORLD, this means that for a world of size n (that is, an $n \times n$ grid) there is a set of 8 actions, n^2 possible positions for the agent, and 2^{n^2} possible configurations of holes. This leads to a set of $1,048,576$ states for a 4×4 TileWorld, and $838,860,800$ in the case of a 5×5 grid. With this rate of growth, the limit for the tractability of direct calculation seems to be at $n = 4$, which is well below what is possible in the BDI model. Even with the many techniques that have been proposed for solving MDPs (for example [4]) that are more efficient than direct calculation, intractability is going to be an issue, and this is one reason why the BDI model is interesting—it can easily handle much larger versions of the TILEWORLD with little problem.

However, these tractability issues don't mean that the MDP model cannot be used at all. The "explosion" in the number of states, as we have seen, depends largely on the amount of holes that can be present at a given moment, and this gives us a means of approximating the solution by pretending that there are fewer holes than there really are. In [17], we present a series of simplifications that allow us to implement MDP agents for a 7×7 TILEWORLD, which permits a more adequate comparison against BDI agents. Current work is being dedicated to obtain performance measures that reflect how good the BDI model's decisions are with respect to those of the MDP agent. Preliminary results show that the approximations to the MDP model outperform BDI; however, extensions to these approximations for a 20×20 grid are in fact outperformed by BDI. Even though further results are required, our initial conclusions are that the BDI model will, in general, outperform the MDP model because the latter will have to resort to simplifications in the face of the complexity inherent to most environments. The strength of the BDI model lies in its use of heuristics, which attacks the complexity of the problem with domain-independent strategies that allow it to make decisions with as much information as possible given the resources that are available.

3 Future Work

As mentioned above, current work is being dedicated to establishing an adequate comparison of the performance of both models in order to gain a better understanding of their relative strengths and weaknesses. We are also currently devoting reasearch in order to establish a relationship between BDI and MDPs. Our goal is to find a way in which the individual components of each model relate to each other, which will enable us to specify one model in terms of the other. Our preliminary research has shown that this is indeed possible, although these early results are primarily of formal interest.

Future work involves establishing formal relationships between both models that allow us to fully understand how the decision making process of one model is seen from the perspective of the other. This is the first step in finding algorithms that, applied to agent specifications in terms of MDPs, obtain specifications in terms of BDI, and *vice versa*. These algorithms would go a long way in bringing the decision-theoretic optimality and ease of specification of MDPs, and the heuristic and intuitive nature of BDI together, aiding the design of autonomous agents by providing the best of both worlds.

References

[1] Adobbati, R., Marshall, A. N., Scholer, A., Tejada, S., Kaminka, G. A., Schaffer, S., and Sollitto, C. Gamebots: A 3d virtual world test-bed for multi-agent research. In *Proceedings of the Second International Workshop on Infrastructure for Agents, MAS, and Scalable MAS* (Montreal, Canada, 2001).

[2] Anderson, J., Baltes, J., Livingston, D., and Sklar, E. Toward an undergraduate league for robocup. In *Proceedings of the RoboCup Symposium* (2003).

[3] Arkin, R. C. *Behavior-based Robotics.* MIT Press, Cambridge, MA, 1998.

[4] Boutilier, C., Dean, T., and Hanks, S. Decision-theoretic planning: Structural assumptions and computational leverage. *Journal of Artificial Intelligence Research 11* (1999), 1–94.

[5] Bratman, M. E., Israel, D., and Pollack, M. Plans and resource-bounded practical reasoning. In *Philosophy and AI: Essays at the Interface*, R. Cummins and J. L. Pollock, Eds. The MIT Press, Cambridge, Massachusetts, 1991, pp. 1–22.

[6] Brooks, R. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation 2*, 1 (1986).

[7] Frias-Martinez, V., Marcinkiewicz, M., Parsons, S., and Sklar, E. The metrobots four-legged league team at robocup 2003. In *Proceedings of RoboCup 2003*, B. Browning and D. Polani, Eds. 2003, p. (to appear).

[8] Georgeff, M. P., and Lansky, A. L. Reactive reasoning and planning. In *Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI-87)* (Seattle, WA, 1987), pp. 677–682.

[9] Kinny, D. N., and Georgeff, M. P. Commitment and effectiveness of situated agents. In *Proc of the Twelfth International Joint Conference on Artificial Intelligence* (San Mateo, California, 1991), B. Grosz and J. Mylopoulos, Eds., Morgan Kaufmann, pp. 82–88.

[10] Pollack, M., and Ringuette, M. Introducing the tileworld: experimentally evaluating agent architectures. In *Proceedings of the Eighth National Conference on Artificial Intelligence* (Menlo Park, CA, 1990), T. Dietterich and W. Swartout, Eds., American Association for Artificial Intelligence, AAAI Press, pp. 183–189.

[11] Puterman, M. L. *Markov decision processes: Discrete Stochastic Dynamic Programming.* John Wiley and Sons, Inc., New York, 1994.

[12] Raiffa., H. *Decision Analysis: Introductory Lectures on Choices under Uncertainty.* Addison-Wesley, 1968.

[13] Schut, M., and Wooldridge, M. Intention reconsideration in complex environments. In *Proceedings of the Fourth International Conference on Autonomous Agents* (Barcelona, June 2000), C. Sierra, M. Gini, and J. S. Rosenschein, Eds., ACM Press, pp. 209–216.

[14] Schut, M., and Wooldridge, M. Principles of intention reconsideration. In *Proceedings of the Fifth International Conference on Autonomous Agents* (Montreal, Canada, May 2001), J. P. Müller, E. Andre, S. Sen, and C. Frasson, Eds., ACM Press, pp. 340–347.

[15] Schut, M., Wooldridge, M., and Parsons, S. Reasoning about intentions in uncertain domains. In *Proceedings of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (Berlin, Germany, 2001), Springer Verlag.

[16] Schut, M. C. *Intention Reconsideration.* PhD thesis, University of Liverpool, 2002.

[17] Simari, G. I., and Parsons, S. D. On the problem of comparing two models for rational decision making in autonomous agents. In *Proceedings of the IX Argentine Congress on Computer Science - CACIC 2003* (La Plata, Buenos Aires, Argentina, 2003), M. Naiouf, Ed., Universidad Nacional de La Plata.