MetroBots Team Description

Michael Littman†, Simon Parsons‡ and Elizabeth Sklar*

†Department of Computer Science
Rutgers University
New Brunswick, NJ, USA

‡Department of Computer and Information Science
Brooklyn College, City University of New York
Brooklyn, NY, USA

*Department of Computer Science
Columbia University
New York, NY, USA.

April 2, 2003

1 Introduction

The MetroBots, a new 4-Legged League team, is a collaboration between[1] Michael Littman at Rutgers University, Simon Parsons at City University of New York (CUNY), and Elizabeth Sklar at Columbia University. The team also includes four Ph.D. students, Paul Batchis (Rutgers), Vanessa Frias-Martinez (Columbia), Dave LeRoux (Rutgers), and Marek Marcinkiewicz (CUNY).

2 Research directions

The aim of the MetroBots project is to advance our research efforts in the area of rational decision making by autonomous agents. There are two main ways in which we are developing this work.

2.1 Decision making and agent models

One major aim we have is to integrate work on decision making under uncertainty, which has led to models such as partially observable Markov decision processes (POMDPs) [5], with work on agent architectures and models for coordinating agents in a multi-agent system.

[1] In alphabetical order.

As things stand, decision making models have largely been developed for rather simple tasks (for example [7]) where the case study is a one-on-one soccer game played in a small grid-world), and the techniques do not scale for real-time, complex applications like the RoboCup task. In contrast, techniques from the multi-agent systems world, like the belief/desire/intention (BDI) model, have been designed to be robust and scaleable, but do not deal well with the uncertainties of interactions in real, physical environments. Our previous work has investigated the use of BDI models in robotics [10] and on the integration between BDI models and POMDPs [13]. The AIBOs seem to be an ideal platform on which to carry out further research on this topic.

The multi-agent systems world has also developed a number of rich models for agent communication and coordination—techniques that can be thought of as mechanisms that allow agents to engage in joint decision making. We have some preliminary results in developing robust techniques for this kind of decision making [11], and aim to use the AIBOs and the soccer task to test these techniques and to drive the further development of them.

2.2 Improving decision making over time

Our second aim we have is to use the AIBOs in developing and testing new approaches to learning behaviors from experience (reinforcement learning and evolutionary learning). The majority of work in reinforcement learning [6] has assumed that decision makers have perfect knowledge about the state of the environment. While this assumption can be approximately satisfied in some domains (like the small robot league with its overhead camera), it ignores some of the most important and interesting aspects of cognition, namely learning and reasoning based on uncertain observations. There is some reinforcement learning work that uses the POMDP framework, which can capture sensor-oriented decision making, but sensors are typically assumed to return fewer than 6 bits of information per time-step.

Little or no work in reinforcement learning has addressed the problem of learning from rich sensor inputs like cameras, and clearly this is the type of learning that will be needed in advanced applications of AI and robotics in the years to come. One concrete task we hope to address early on is creating players that can develop "field sense" by comparing situations with a large corpus of experience and using it to judge which decisions will be most effective. We feel that that AIBO platform and the RoboCup task are ideally suited to this kind of research.

Both evolutionary and reinforcement learning are traditionally difficult on real robots because the number of iterations it takes for a learning algorithm to converge is typically longer than the battery life of a real robot. In addition, evolutionary learning in simulation has traditionally been a problem for robotics, because the behaviours learned in simulation do not transfer well onto robots due to the uncertainties in the environment which the simulator does not model. Evolutionary learning has the additional problem of needing sometimes large populations of agents to learn from, and these cannot be run on an onboard processor. We plan to take advantage of the wireless connection with the AIBO to feed real-time learning engines that run in parallel with the AIBO. As the AIBO experiences the world, it sends environmental parameters to the learning engines. As the learning engines progress, they will send improved behaviours to the AIBO.

2.3 Approach to RoboCup challenge

To date we have only had the AIBOs for a few months. We are working to obtain robust vision and localization, and basic behaviours (such as scanning for markers, moving to the ball, and kicking towards the opponents goal). Once these basics are complete, we will start to address the research issues described above.

At the centre of our approach is the use of a probabilistic (Monte-Carlo) localization. This builds a probabilistic map of the pitch which integrates in a natural way with a POMDP decision making model (in which the state transitions are achieved by the simple behaviours). This representation also forms the basis of the kind of learning models that we will use. POMDPs are essentially a reinforcement learning model, and the learning will help to refine the state-transition probabilities and the probability distributions used to update with sensor readings. Evolutionary computation will also be used to refine these parameters. Once we have a suitably accurate model, we can abstract from this to create a BDI model that retains the essential features of the full POMDP, but is computationally tractable.

3 Background of investigators

The team leaders are experienced researchers in the field of artificial intelligence.

Michael Littman's research is in the area of machine learning, examining algorithms for decision making under uncertainty and statistical natural language processing. After earning his Ph.D. from Brown University in 1996, Michael worked as an assistant professor at Duke University, a member of technical staff in AT&T's Artificial Intelligence Principles Research Department, and is now an associate research professor of computer science at Rutgers. He is an associate editor of the Journal of Artificial Intelligence Research and an editorial board member of the Journal of Machine Learning Research.

Simon Parsons' research centres around the problem of coordinating action in multi-agent systems. Solving this problem involves making progress in inter-agent communication, learning and decision making under uncertainty, and these are all areas in which he has made research contributions. After obtaining his PhD from the University of London in 1993, Simon worked as a faculty member at the Universities of London and Liverpool until December 2001. For the first nine months of 2002 he was a visiting scholar at MIT, and since September 2002 has been at Brooklyn College, where he is an associate professor. Simon has been editor of Knowledge Engineering Review since 1998.

Elizabeth Sklar's research revolves around evoluationary learning in agents. She has focused on learning in populations of agents within simulated gaming environments (e.g., robot hockey and simple video games). Much of her work on learning has involved integrating usage data, gathered from humans over the internet, into a real-time learning engine. Sklar obtained her PhD in 2000 from Brandeis University. Subsequently, she was a visiting assistant professor at Boston College for one year and since then has been an assistant professor at Columbia University. She was the chair of RoboCupJunior 2001 and RoboCupJunior 2002. She is currently chair of the RoboCupJunior International Technical Committee and is a member of the Executive Committee of RoboCup. She is also the co-chair of the Java Engagement for Teacher Training project, organized through the ACM to improve secondary computer science education.

4 Team organization

As a result of the geographic distribution of the team members, it is not possible for us to work as one unit. The New York groups (the group at Columbia and the group at Brooklyn College) work closely together, meeting weekly and sharing lab space. The New Jersey (Rutgers) group work together on a daily basis. The whole team (from all three universities) meet every three weeks or so..

As a new team, we have been struggling, as one might expect, with acquiriing equipment and the basic tasks of vision and localization, and this has so far consumed all of our time. The four students mentioned above have been working, and will continue to work, full time on the project. In the future, once the low level modules are working efficiently and robustly, we anticipate that other students, both graduate and undergraduate, will be able to help. Indeed our goal is to use the AIBOs as platforms to advance all of our major research interests (as detailed above) and, as a result, the team will benefit from the involvement of all of our students.

5 Relevant publications

As a team that is new to the competition, we have no publications about our work on MetroBots. However, we bring to this work our considerable experience of related research, including:

- Reinforcement learning: [6], [8], [14].
- Evolutionary learning: [1], [3], [4], [12].
- Agent decision making: [5], [7], [10], [13], [16].
- Simulated agent-based gaming environments: [2], [15].
- Agent communication and coordination: [9], [11].

6 Commitment to attend

Three members of the MetroBots team (Frias-Martinez, Parsons and Sklar) will attend RoboCup whether or not the team qualifies (they have already purchased airline tickets and will be attending as part of the organisation of RoboCupJunior), and they will be joined by Batchis, LeRoux and Marcinkiewicz if the team qualifies. The team has funds to cover registration.

References

[1] A. Blair and E. Sklar. The evolution of subtle manoeuvres in simulated hockey. In et al. R. Pfeifer, editor, *Proceedings Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior (SAB-98)*. MIT Press, 1998.

[2] A. Blair and E. Sklar. Exploring evolutionary learning in a simulated hockey environment. In *Proceedings of the 1999 Congress on Evolutionary Computation (CEC-99)*, 1999.

[3] A. Blair, E. Sklar, and P. Funes. Co-evolution, Determinism and Robustness. In *Simulated Evolution and Learning (SEAL-98), Lecture Notes in Artificial Intelligence 1585*, pages 389–396. Springer-Verlag, 1998.

[4] P. Funes, E. Sklar, H. Juillé, and J. Pollack. Animal-Animat Coevolution: Using the Animal Population as Fitness Function. In et al. R. Pfeifer, editor, *Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior (SAB-98)*. MIT Press, 1998.

[5] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998.

[6] Leslie Pack Kaelbling, Michael L. Littman, and Andrew W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[7] Michael L. Littman. Value-function reinforcement learning in Markov games. *Cognitive Systems Research*, 2(1):55–66, 2001.

[8] Stephen M. Majercik and Michael L. Littman. Contingent planning under uncertainty via stochastic satisfiability. *Artificial Intelligence*, 2002. In press.

[9] S. Parsons and N. R. Jennings. Negotiation through argumentation — a preliminary report. In *Proceedings of Second International Conference on Multi-Agent Systems*, pages 267–274, 1996.

[10] S. Parsons, O. Pettersson, A. Saffiotti, and M. Wooldridge. Intention reconsideration in theory and practice. In W. Horn, editor, *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI-2000)*. John Wiley, 2000.

[11] S. Parsons, M. Wooldridge, and L. Amgoud. Properties and complexity of formal inter-agent dialogues. *Journal of Logic and Computation*, (to appear), 2003.

[12] S. Phelps, P. Mcburney, S. Parsons, and E. Sklar. Co-evolutionary mechanism design: A preliminary report. In J. Padget, D. C. Parkes, N. M. Sadeh, O. Shehory, and W. E. Walsh, editors, *Agent-Mediated Electronic Commerce IV*. Spinger Verlag, Berlin, Germany, 2002.

[13] M. C. Schut, M. Wooldridge, and S. Parsons. Reasoning about intentions in uncertain domains. In D. Dubois and H. Prade, editors, *Proceedings of European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Toulouse, France, 2001.

[14] Satinder Singh, Tommi Jaakkola, Michael L. Littman, and Csaba Szepesvári. Convergence results for single-step on-policy reinforcement-learning algorithms. *Machine Learning*, 39:287–308, 2000.

[15] E. Sklar, A. Blair, and J. Pollack. Training intelligent agents using human data collected on the internet. In *Agent Engineering*. World Scientific, Singapore, 2001.

[16] M. Wooldridge and S. Parsons. Intention reconsideration reconsidered. In J. P. Müller, M. P. Singh, and A. Rao, editors, *Intelligent Agents V*, pages 63–80. Springer Verlag, 1999.