Final Summary of Research

Self-Directed Cooperative Planetary Rovers

NASA Grant Number NAG 2-1463

Shlomo Zilberstein, P.I.
Department of Computer Science
University of Massachusetts at Amherst
Amherst, MA 01003

April 2003

1 Overview

This final summary of research describes work done under NASA Grant Number NAG 2-1463 during the period 3/1/2001–2/28/2003. Note that this project has been converted to a Cooperative Agreement Number NCC 2-1311 after the first year of the project. The remaining work on the project is being conducted under this cooperative agreement.

The project is concerned with the development of decision-theoretic techniques to optimize the scientific return of planetary rovers. Planetary rovers are small unmanned vehicles equipped with cameras and a variety of sensors used for scientific experiments. They must operate under tight constraints over such resources as operation time, power, storage capacity, and communication bandwidth. Moreover, the limited computational resources of the rover limit the complexity of on-line planning and scheduling. We have developed a comprehensive solution to this problem that involves high-level tools to describe a mission; a compiler that maps a mission description and additional probabilistic models of the components of the rover into a Markov decision problem; and algorithms for solving the rover control problem that are sensitive to the limited computational resources and high-level of uncertainty in this domain.

The project is directed by Shlomo Zilberstein at the University of Massachusetts in collaboration with Eric Hansen (Mississippi State Univ.), Victor Lesser (Univ. of Massachusetts) and Richard Washington (NASA Ames). In addition to the co-investigators, five graduate students, a post-doctoral research fellow, and two unfunded collaborators from France have worked on the project

during the period covered by the report.

2 Summary of Research Accomplishments

This section summarizes the main accomplishments of this project. These accomplishments are described in detail in the attached publications.

2.1 Modeling a mission as a sequential decision process

One of the fundamental premises of our work is the ability to translate the rover's mission into a sequential decision process. We have gathered information on the scope of rover operation in missions planned for the coming decade and identified the type of planning constraints that must be captured in order to express the range of activities in these missions. We have developed a language for high-level modeling of rover missions and software tools for automatically generating a corresponding Markov decision process (MDP) that represents the rover control problem. Finally, we built a partial probabilistic model of the K9 rover used at NASA Ames. These results enable us to generate a variety of realistic mission plans for testing and evaluation of our algorithms. As part of this effort, Max Horstmann, a UMass graduate student, participated in the NASA Ames/RIACS 2002 Summer Student Research Program. He worked with Rich Washington, Nico Meuleau, and others on probabilistic modeling of rover missions.

2.2 Solving large MDPs using hierarchical reinforcement learning

Weakly-coupled Markov decision processes can be decomposed into subprocesses that interact only through a small set of bottleneck states. The MDPs that capture the mission of an autonomous rover are weakly-coupled, because there is limited interaction between the scientific data collected in the course of the mission. We have studied a hierarchical reinforcement learning algorithm designed to take advantage of this particular type of decomposability. The algorithm was tested using a simple simulator of an autonomous planetary rover. In our tests, a Mars rover must decide which activities to perform and when to traverse between sites in order to make the best use of its limited resources. In our experiments, the hierarchical algorithm performed better than Q-learning in the early stages of learning, but unlike Q-learning it converged to a suboptimal policy. This suggests that it is advantageous to use the hierarchical algorithm when training time is limited. We continue to study this novel algorithm for hierarchical reinforcement learning. An ECP-2001 paper descibing this work is attached.

2.3 Solving large MDPs using adaptive decision-theoretic planning

We have developed another decision-theoretic framework for optimizing the scientific return of planetary rovers. This framework is based on giving the rover multiple methods in which to accomplish each step of a plan. The different alternatives offer a tradeoff between resource consumption and the quality of the outcome. We have shown how to choose the best way to execute a task based on the availability of resources, the progress made with the task so far, and the remaining workload. Each task is controlled by a precompiled policy that factors the effect of the remaining plan using the notion of an opportunity cost. An attached paper describing this work was presented at the

Dagstuhl Workshop on Plan-based Control of Robotic Agents in 2001 and was published in LNAI No. 2466 in 2002.

2.4 Symbolic heuristic search for factored MDPs

The complexity of the planning algorithms we previously developed for planetary rovers can be significantly reduced by exploiting the structure of the domain and using admissible heuristic search. To address the potentially large size of the state space, we have developed algorithms that use state abstraction to avoid evaluating states individually. Forward search from a start state, guided by admissible heuristic, is used to avoid evaluating all states. The two have been combined in a novel way that exploits symbolic model-checking techniques. This work, conducted by Zhengzhu Feng and Eric Hansen, was presented at AAAI-2002. The paper is attached to the report.

2.5 Generating understandable contingency plans for rovers

The planning algorithms we have developed for planetary rovers produce a plan represented as a policy that maps states to actions. Such policies are provably optimal with respect to the probabilistic model we use, but they are not easy to understand or analyze. They could also require a large amount of storage. To address these weaknesses, Max Horstmann completed an MS project aimed at converting MDP policies to much more understandable contingency plans. He developed a general representation of contingency plans, a numerical measure of clarity, and algorithms for optimizing the clarity of the plan. Clarity is a measure of the compactness of the plan in terms of the number of nodes and branches. When the model used for constructing the plan is approximate, the contingency plan representation helps to reveal counter intuitive or undesirable patterns in the rover control plan. Such patterns are much harder to detect using a "flat" policy representation. This work was highly influenced by the interactions that Max had during his SSRP internship in 2002 with several researchers at NASA Ames.

2.6 Control of multiple rovers

Control of multiple rovers can be modeled as a form of decentralized Markov decision process. We analyzed the complexity of decentralized MDPs and showed that the problem is NEXP-hard. This means that deriving optimal plans for two or more cooperating rovers is extremely difficult. We have made substantial progress with this problem. We identified a class of problems called *transition-independent* MDPs, that captures effectively the control problem of multiple rovers. The general class consists of independent collaborating agents that are tied up by a global reward function that depends on both of their execution histories. For example, when two rovers are deployed, each with its own mission, there is important interactions between the activities they perform. The activities may be *complementary* (e.g., taking pictures of two sides of a rock), or the y may be *redundant* (e.g., taking two spectrometer readings of the same rock). We developed a novel algorithm for solving this class of problems and examined its properties. This result is the first effective techniques to solve optimally a class of decentralized MDPs. This work, conducted by Raphen Becker, Shlomo Zilberstein, Victor Lesser, and Claudia Goldman, will be presented at AAMAS 2003. This paper is attached to the report.

2.7 Interface with the experimental platforms used at NASA Ames

The focus of this project is on high-level, decision-theoretic rover control. This rests on a number of existing layers of control, which bridge the gap from decision-theoretic plans to the low-level control of the robotic mechanisms. For evaluation purposes, we are targeting our work for the NASA Ames "K9" rover prototype. The existing rover software architecture consists of four distinct layers. Low-level device drivers communicate with hardware. Mid-level component controllers receive simple commands (such as direct movement, imaging, and instrument commands) and communicate with the device drivers to effectuate the commands. Abstract commands implement compound or complex actions (such as movement with obstacle avoidance, visual servoing to a target, and arm placement). A plan executive interprets command plans and calls both simple and abstract commands as specified in the plan. We have designed our high-level, decision-theoretic controller to interact with this architecture by decomposing actions into subplans; these are provided to the rover plan executive, which in turn manages the execution and monitoring of the subplans. Information about action success and the resulting state of the system is returned to the decision-theoretic controller.

To facilitate this integration, Max Horstmann spent much of the summer of 2002 at NASA Ames experimenting with the K9 platform. Building on earlier work and visits by another student, Dan Bernstein, Max confirmed the feasibility of our approach which relies on translating fragments of the high-level plan into CRL (the Contingent Rover Language developed at NASA) and passing them to the rover plan executive. We anticipate to continue with similar activities in the future to guarantee the compatibility of this effort with the experimental platforms used by NASA and to simplify future technology transfer.

3 Publications

1. D.S. Bernstein, R. Givan, N. Immerman, and S. Zilberstein. "The Complexity of Decentralized Control of Markov Decision Processes." *Mathematics of Operations Research*, Vol. 27, No. 4, 819–840, November 2002.

2. S. Zilberstein, R. Washington, D.S. Bernstein, and A.I. Mouaddib. "Decision-Theoretic Control of Planetary Rovers." In M. Beetz et al. (Eds.): Plan-Based control of Robotic Agents, *LNAI*, No. 2466, 270-289, November 2002. **(Attached)**

3. M. Horstmann. "Automated Generation of Understandable Contingency Plans for Markov Decision Processes." MS Project Report, University of Massachusetts, December 2002.

4. C.V. Goldman and S. Zilberstein. "Optimizing Information Exchange in Cooperative multi-agent Systems." To appear in *Second International Joint Conference on Autonomous Agents and Multi-Agent Systems*, July 2003.

5. R. Becker, S. Zilberstein, V. Lesser, and C.V. Goldman. "Transition-Independent Decentralized Markov Decision Processes." To appear in *Second International Joint Conference on Autonomous Agents and Multi-Agent Systems*, July 2003. **(Attached)**

6. E. Hansen, R. Zhou and Z. Feng. "Symbolic Heuristic Search using Decision Diagrams." *Fifth International Symposium on Abstraction, Reformulation and Approximation (SARA-02)*, Edmonton, Alberta, Canada, July 2002.

7. Z. Feng and E. Hansen. "Symbolic Heuristic Search for Factored Markov Decision Processes." *Proceedings of the Eighteenth National Conference on Artificial Intelligence*, Edmonton, Alberta, Canada, July 2002. **(Attached)**

8. Z. Feng and E. Hansen. "Symbolic LAO* Search for Factored Markov Decision Processes." *AIPS 2002 Workshop on Planning via Model Checking*, Toulouse, France, April 2002.

9. D. S. Bernstein and S. Zilberstein. "Reinforcement Learning for Weakly-Coupled MDPs and Application to Planetary Rover Control." *Proceedings of the Sixth European Conference on Planning*, Toledo, Spain, September, 2001. **(Attached)**

10. Z. Feng and E.A. Hansen. "Approximate Planning for Factored POMDPs." *Proceedings of the Sixth European Conference on Planning*, Toledo, Spain, September, 2001.

11. S. Cardon, A. I. Mouaddib, S. Zilberstein, and R. Washington. "Adaptive Control of Acyclic Progressive Processing Task Structures." *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, pp. 701–706, Seattle, Washington, August, 2001.

12. D.S. Bernstein, S. Zilberstein, R. Washington, and J. Bresina. "Planetary Rover Control as a Markov Decision Process." *Proceedings of the Sixth International Symposium on Artificial Intelligence, Robotics and Automation in Space*, Montreal, Canada, June, 2001.

13. D. S. Bernstein and S. Zilberstein. "Reinforcement Learning for Weakly-Coupled MDPs and an Application to Planetary Rover Control." *ICML Workshop on Hierarchy and Memory in Reinforcement Learning*, Williamstown, Massachusetts, June, 2001.

14. P. Xuan, V. Lesser, and S. Zilberstein. "Communication Decisions in Multi-agent Cooperation: Model and Experiments." *Proceedings of the Fifth International conference on Autonomous Agents*, pp 616–623, Montreal, Canada, June, 2001.

15. S. Zilberstein and A. I. Mouaddib. "Adaptive Planning and Scheduling of On-board Scientific Experiments." *European Space Agency (ESA) Workshop on On-Board Autonomy*, Noordwijk, The Netherlands, October, 2001.

16. D. S. Bernstein, T. J. Perkins, S. Zilberstein, and L. Finkelstein. "Scheduling Contract Algorithms on Multiple Processors." *AAAI Fall Symposium on Using Uncertainty within Computation*, North Falmouth, Massachusetts, November, 2001.

17. D. S. Bernstein and S. Zilberstein. "Planetary Rover Control as a Markov Decision Process." *AAAI Spring Symposium on Game-Theoretic and Decision-Theoretic Agents*, Stanford, California, March, 2001.

4   Project Related Invited Talks

1. The PI gave a keynote presentation entitled "Decision-Theoretic Control of Autonomous Planetary Rovers" at the *Dagstuhl Seminar on Plan-Based Control of Robotic Agents*, Dagstuhl, Germany, October, 2001.

5 Project Related Panel Discussions

1. The PI served as a member of a panel entitled "Integrated System Architecture for On-board Autonomy" at the *European Space Agency Workshop on On-Board Autonomy*, Noordwijk, The Netherlands, October, 2001.