OxfordRescue Team 2008
RoboCup Rescue - Virtual Robots Competition

Julian de Hoog

Computing Laboratory
University of Oxford
Parks Road, Oxford OX1 3QD
United Kingdom
julian.dehoog@comlab.ox.ac.uk

Abstract. OxfordRescue will participate in the 2008 Virtual Robots competition. In addition to building on successful approaches employed by other teams in such areas as mapping, victim detection and practical user interfaces, OxfordRescue intends to apply reinforcement learning techniques to problems that arise in communication and cooperation. Since search and rescue scenarios often suffer from communication limitations, a "Messenger" approach will be applied to the multi-robot exploration task that dynamically assigns different roles to robots, creating an elastic communication network. It is hoped that this approach will increase the efficiency of exploration by allowing certain robots to explore without interruption, while others relay information back and forth through the team. The network is represented using a context-specific coordination graph and individual action payoffs are to be learnt using Q-learning.

1 Introduction

The OxfordRescue Team will represent the University of Oxford at the 2008 RoboCup Rescue event, in the Virtual Robots Competition. The main challenges of the Virtual Robots competition are typically identified as the following: Navigation, Mapping, Victim Detection, Communication, Cooperation and Practical Interface Development [1, 2]. Numerous advances have already been made by various teams in each of these categories.

OxfordRescue hopes to build on these teams' successes and incorporate some of their many positive approaches within a single unified team. For example, it could be of great benefit to use a mapping approach similar to the University of Amsterdam's successful MainfoldSLAM method [3, 2], while using a human-robot interface approach similar to the one pioneered by the University of Pittsburgh [4]. Moreover, however, OxfordRescue hopes to move into a new direction by applying machine learning techniques to the problems of communication and cooperation.

Robots operating in real search and rescue scenarios are likely to experience severe communication limitations. The importance of this real-life constraint is reflected in the organizers' decision to introduce limitations on communication from the 2007 competition onwards. For a team of robots exploring a disaster zone, it can be useful to share the information acquired by individual robots. This is important not only for development of a single unified map, but also for an increased efficiency of the exploration effort as a whole. OxfordRescue hopes to apply machine learning techniques to this problem of "Distributed Information Management", and thereby contribute to the RoboCup Rescue community.

2 Machine Learning

Machine learning techniques can provide robust solutions to problems in which agents may learn useful behavior as a result of past experience. Numerous techniques have been applied with success to multi-agent systems, and many techniques have provided successful solutions even in noisy and uncertain environments.

In RoboCup, machine learning has been applied with success to numerous other competitions. For example, the current champion in the 2D soccer simulation is the University of Osnabrueck's Brainstormers, which is based to a large extent on reinforcement learning principles [5]. Reinforcement learning has also been applied to numerous subtasks of this competition, e.g. [6, 7]. Latzke *et al.* have used a modified type of reinforcement learning to train a humanoid robot to learn fundamental soccer skills [8]. Evolutionary reinforcement learning has also been applied to the Agents Simulation with success by Martínez *et al.* [9].

The Virtual Robots competition shares a number of properties with RoboCup's other competitions: a very large state space, multiple agents acting independently, hidden and uncertain information, limited communication and trade-offs between short-term goals and long-term goals. Since learning techniques have proven successful in RoboCup's other leagues, it is likely that they could be applied with success to the Virtual Robots competition.

3 Distributed Information Management

Though multi-robot exploration is a relatively young field, a number of different approaches have demonstrated success [10–13]. These approaches all rely on, or are improved by, joint communication between all members of the robot team. As has been discussed, search and rescue scenarios frequently suffer from severe communication limitations. In such environments, it is inefficient for all robots to try to stay within communication range of all other robots. It makes much more sense for robots to explore the far reaches of the environment and to use another mechanism to pool the information acquired by each. Possible solutions to this problem include:

1. Making sure that all robots stay within a team-wide multi-hop communication network. For example, while robots A and B at far reaches of the environment may not be able to communicate directly, they may relay information through robots C and D between them that are within range of each of them respectively, and of one another. In a sense, this approach has been implemented in [14], where individual robots use a bidding algorithm to explore that makes robots stay close to one another.
2. Letting individual robots explore on their own, but arranging for them to return for a "rendezvous" at a specific location and time to pool their knowledge, as suggested in [15]. In the Virtual Robots competition, this is equivalent to robots returning to the ComStation just before the end of the 20-minute time limit.
3. Using a "Messenger" approach: allow most of the robots in the team to undertake exploration, but also use a small number of robots to relay information back and forth between robots, or groups of robots, that are out of range. Ideally, the number of messengers would grow and shrink depending on the extent of the exploration effort.

A possible drawback to method 1 is inefficiency: while communication of data is quick and efficient, the team of robots must navigate the environment as a sort of "cloud" that must stay together. For large environments, there is a strong possibility of terrain being left unexplored. Method 2 is likely to lead to more efficient exploration, but likely also leads to time being wasted when robots must return to the agreed meeting point. This drawback is accounted for by method 3. While the "messenger" robots are not contributing to the exploration effort, the "explorer" robots at the far reaches of the environment may continue to explore without having to waste time returning to an agreed upon meeting point. Moreover, the pool of joint team knowledge is updated regularly. It could be of great interest to develop such an approach and evaluate it within the Virtual Robots framework.

The team network as a whole could expand dynamically, with information being passed from far reaches of the environment along a chain of messengers. Ideally robots would be assigned roles dynamically as the team network expands, with explorers becoming messengers as required.

Elsewhere, multi-agent cooperation has been achieved using context-specific coordination graphs ([16, 17]). A similar approach could be used here: each node in the graph represents one of the robots in the exploration team. Edges are dynamically created between robots having recently had a communication link and removed between robots that are out of range. Each robot has a specific role, such as Explorer or Messenger. Each role has a corresponding set of possible actions that may be pursued (e.g. explore immediate vicinity, move deeper into the environment, relay information, etc.), represented as value rules and having corresponding payoffs. These payoffs depend on nearby robots (for example if two robots are next to one another, it makes sense for one to explore nearby rooms and the other to move deeper into the environment). A maximum global payoff, repeatedly computed online, determines individual robots' actions, and possibly changes in roles.

As has proven successful elsewhere, individual action payoff values could be learnt using reinforcement learning, in particular Q-learning. The joint action space would become large quickly, but various techniques such as variable elimination [18], the max-plus algorithm [19] and simulated annealing [20] have proven efficient in finding an optimal or near-optimal global solution.

4 System Architecture

Given the nature of OxfordRescue's approach, it will likely make sense to use a heterogeneous team of robots in the exploration effort. Explorers should be well-equipped to locate victims, while messengers should ideally be quick. A likely combination could include robust P2AT robots together with smaller and faster P2DX robots.

For outdoor scenarios, it could be of great interest to use ATRV-Jr's. In the virtual robots competition, emphasis is justifiably placed on real-world applicability, as demonstrated by various studies of validation of the simulator [21, 22]. The use of ATRV-Jr's leaves open the door for a possible collaboration with Oxford's Department of Engineering on a potential future validation study.

5 Conclusion

Developing a successful team in the Virtual Robots competition requires solutions to a wide range of problems. It is not feasible for an individual team to make significant advances in every area of the competition. As a result, OxfordRescue hopes to build on the many highly interesting and effective approaches adopted by other teams in the Virtual Robots competition in such areas as mapping, victim detection, and practical interface development. However, OxfordRescue hopes to additionally conduct research into the possible application of machine learning techniques to further problems in communication and cooperation, including the management of distributed information given communication limitations. Hopefully the application of some of the ideas presented here will demonstrate a degree of improved performance and will lead to a useful contribution to the RoboCup Rescue community.

References

1. Balakirsky, S., Carpin, S., Kleiner, A., Lewis, M., Visser, A., Wang, J., Ziparo, V.: Towards heterogeneous robot teams for disaster mitigation: Results and performance metrics from robocup rescue. Journal of Field Robotics **24** (2007) 943 – 967

2. Visser, A., Slamet, B., Schmits, T., Jaime, L.A.G., Ethembabaoglu, A.: Design decisions of the UvA Rescue 2007 Team on the challenges of the virtual robot competition. In: Proceedings of the Fourth International Workshop on Synthetic Simulation and Robotics to Mitigate Earthquake Disaster. (2007) 20 – 26
3. Slamet, B., Pfingsthorn, M.: ManifoldSLAM: a multi-agent simultaneous localization and mapping system for the robocup rescue virtual robots competition. Master's thesis, Universiteit van Amsterdam (December 2006)
4. Wang, J., Lewis, M.: Human control for cooperating robot teams. In: HRI '07: Proceedings of the ACM/IEEE international conference on Human-robot interaction, New York, NY, USA, ACM (2007) 9 – 16
5. Riedmiller, M., Gabel, T.: Brainstormers 2d - team description 2007. In Visser, U., Ribeiro, F., Ohashi, T., Dellaert, F., eds.: RoboCup 2007: Robot Soccer World Cup XI, LNCS, Springer (2007)
6. Kalyanakrishnan, S., Liu, Y., Stone, P.: Half field offense in robocup soccer: A multiagent reinforcement learning case study. In: RoboCup 2006: Robot Soccer World Cup X. (2007) 72 – 85
7. Stone, P., Sutton, R.S., Kuhlmann, G.: Reinforcement learning for RoboCup-soccer keepaway. Adaptive Behavior **13**(3) (2005) 165–188
8. Latzke, T., Behnke, S., Bennewitz, M.: Imitative reinforcement learning for soccer playing robots. In: RoboCup 2006: Robot Soccer World Cup X. (2007) 47 – 58
9. Martínez, I.C., Ojeda, D., Zamora, E.A.: Ambulance decision support using evolutionary reinforcement learning in robocup rescue simulation league. In: RoboCup 2006: Robot Soccer World Cup X. (2007) 556 – 563
10. Yamauchi, B.: Frontier-based exploration using multiple robots. In: AGENTS '98: Proceedings of the second international conference on Autonomous agents, New York, NY, USA, ACM (1998) 47 – 53
11. Simmons, R.G., Apfelbaum, D., Burgard, W., Fox, D., Moors, M., Thrun, S., Younes, H.: Coordination for multi-robot exploration and mapping. In: AAAI/IAAI. (2000) 852–858
12. Zlot, R., Stentz, A., Dias, M., Thayer, S.: Multi-robot exploration controlled by a market economy. Robotics and Automation, 2002. Proceedings. ICRA '02. IEEE International Conference on **3** (2002) 3016 – 3023
13. Burgard, W., Moors, M., Stachniss, C., Schneider, F.: Coordinated multi-robot exploration. IEEE Transactions on Robotics **21**(3) (2005) 376 – 378
14. Sheng, W., Yang, Q., Tan, J., Xi, N.: Distributed multi-robot coordination in area exploration. Robot. Auton. Syst. **54**(12) (2006) 945 – 955
15. Roy, N., Dudek, G.: Collaborative robot exploration and rendezvous: Algorithms, performance bounds and observations. Journal of Autonomous Robots **11**(2) (2001) 117 – 136
16. Guestrin, C., Venkataraman, S., Koller, D.: Context-specific multiagent coordination and planning with factored mdps. In: Eighteenth national conference on Artificial intelligence, Menlo Park, CA, USA, American Association for Artificial Intelligence (2002) 253 – 259
17. Kok, J.R., Spaan, M.T.J., Vlassis, N.: Multi-robot decision making using coordination graphs. In de Almeida, A.T., Nunes, U., eds.: Proceedings of the International Conference on Advanced Robotics (ICAR), Coimbra, Portugal (June 2003) 1124 – 1129
18. Kok, J.R., Vlassis, N.: Multiagent Q-learning by context-specific coordination graphs. In Groen, F., Amato, N., Bonarini, A., Yoshida, E., Kröse, B., eds.: Proceedings of the International Conference on Intelligent Autonomous Systems, Amsterdam, The Netherlands, IOS Press (March 2004) 317 – 324

19. Kok, J.R., Vlassis, N.: Using the max-plus algorithm for multiagent decision making in coordination graphs. In: RoboCup-2005: Robot Soccer World Cup IX, Osaka, Japan (July 2005)
20. Dawei, J., Shiyuan, W.: Using the simulated annealing algorithm for multiagent decision making. In: RoboCup 2006: Robot Soccer World Cup X. (2007) 110 – 121
21. Wang, J., Lewis, M., Hughes, S., Koes, M., Carpin, S.: Validating USARSim for use in HRI research. In: Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting. (2005) 457 – 461
22. Carpin, S., Stoyanov, T., Nevatia, Y., Lewis, M., Wang, J.: Quantitative assessments of USARSim accuracy. In: Proceedings of PerMIS