UvA Rescue
Team Description Paper
Multi-Agent Challenge
Rescue Simulation League
RoboCup 2014 - João Pessoa - Brazil

Mircea Traichioiu and Arnoud Visser

Universiteit van Amsterdam, Science Park 904, 1098 XH Amsterdam, NL
http://www.jointrescueforces.eu

Abstract. The UvA Rescue Team likes to concentrate their effort inside the Rescue Simulation League to a theoretical contribution, which is accordance with the Multi-Agent Challenge. The UvA Rescue Team has formulated the planning and coordination problem formally as a POMDP problem, which will allow to apply POMDP-solution methods in this application area. Inside the Multi-Agent Challenge the limits on communication no longer applies, which allows us to model communication actions as instantaneous and not limited by distance or bandwidth.

1 Introduction

The UvA Rescue Team has a long history. Yet, this will be the first participation inside the Multi-Agent Challenge. The UvA Rescue team has already experience in the Agent competition [8, 12] and the Virtual Robot competition was held [7, 13].

During those years the team published several theses, conference papers and journal articles. The latest article [1] gives a nice overview of the RoboCup Rescue Leagues.

2 Context

The Multi-Agent Challenge decomposes the problem of coordinating a large team of emergency responders into several subproblems, as described in [2].

The first subproblem is assigning fire-brigade agents to firefronts. This is a real multi-agent problem, working together could have non-linear benefits. In addition, too many agents can be counterproductive. The fire-agents for this benchmark could choose between three actions [3], while our fire-agents inside the Agent competition already have to consider seven actions [12].

This reduction in action space will allow to consider larger teams or to enlarge the planning horizon. In [2] several coordination algorithms were tested for the city Kobe with three ignition points, 12 fire-brigade agents and simulating 300 timesteps.

The second subproblem is to support the evacuation of 3000 civilians from a city (Berlin or Paris) after an earthquake. Many roads are blocked by debris, which could be cleared by police agents.

3 Approach

The intention of our team is to formulate the coordination problem of the Agent competition in such a way that solution methods of the MultiAgent Decision Process toolbox [9] can be used. The toolbox expects the coordination problem to be described as Distributed Partial Observable Markov Decision Process (DEC-POMDP), which means that the effect of the actions is stochastic and the state of the world is only partial observable. The benchmark describes the coordination as a Distributed Constraints Optimization Problem (DCOP), which means that the effect of the actions is deterministic and only the parameters of actions have to be optimized. Both problem descriptions assume that the reward function is known. In contrast, Distributed Coordination of Exploration and Exploitation (DCEE) agents can also reason about uncertain rewards (but not about the uncertainty in the action outcomes) [11].

Under the general DCOP framework, each agent is assigned a variable, the goal being to determine an optimal value for that variable that maximizes some global utility measure. This utility is defined through an utility matrix specifying values for different joint configurations of variable values. In particular for the firefighting scenario, the utility matrix described in the RMASBench framework is a bidimensional matrix $U_{a,t}$ reflecting reflecting the utility of agent a attending burning building (target) t. This measure is computed based on the fieriness of the burning building and distance between the agent and the building. Thus, for this particular problem, the goal for the DCOP solver is to compute the optimal assignment of fire agents to burning buildings. At each time step the procedure is repeated, the reformulation of the "problem" taking into account the changes in the environment.

In contrast to the DEC-POMDP methods used to address the firefighting and victim rescuing (i.e. addressing competition goals directly), described in [12], solving the DCOP problem has a different goal, namely the optimal assignment of tasks at each time step, which only indirectly leads to an efficient firefighting and victim rescuing. Moreover, reasoning about the DCOP problem happens once every simulation step, and not throughout the entire simulation episode. Thus, the model for this problem can be formulated as a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \Omega, \mathcal{O}, T, p_0 \rangle$ where:

- $\mathcal{S}$ represents the state space, each state denoting a joint assignment configuration;
- $\mathcal{A}$ represents the action space, each agent's action altering its own assignment choice;

- $\mathcal{P}(s, a, s')$ represents the transition funtion, reflecting probability of arriving in state s' from state s, taking action a; Although obviously each agent will deterministically change his assignment choice through his action a, it cannot be certain regarding the other agent's actions, whose results are still encoded in the state s';
- $\mathcal{R}(s)$ is a reward function computed based on the known utility matrix and the assignment configuration described by the state s;
- Ω represents the set of observations, which reflect the knowledge of the agent regarding the state it is currently in;
- $\mathcal{O}(s, a, o_1...o_n, s')$ is a function reflecting the observation probabilities;
- T represents the time horizon of the problem;
- p_0 represents the start state distribution.

The partial observability property of this problem arises from the fact that in general an agent cannot be aware of all other agent's action choices. However, in a more particular case of perfect communication, agents would be able to broadcast their own choices to all other agents and thus each of them would be able to know for certain the state in which they are in. Even if this eliminates the partial observability, the stochasticity of the actions still remains, as the agents choose their actions simultaneously and are unable to exactly predict the state in which they will transition to.

The time horizon of the problem reflects the estimated running time constraints of each simulation step and insures a sufficiently good solution will be computed before the allocated time expires.

Having this formal definition of the problem, standard planning algorithms available in the MADP toolbox can be applied.

4 The MultiAgent Decision Process Toolbox

The MADP toolbox[1] contains several planning algorithms. The planning library depends on the other contained libraries and offers functionality for the planning algorithms. In particular the library contains:

- *Dec-POMDP solution algorithm:* BruteForceSearchPlanner, JESPExhaustivePlanner, JESPDynamicProgrammingPlanner [4], DICEPSPlanner [5], k-GMAA* and GMAA* [6].
- *POMDP solution techniques:* Perseus [10].
- *Heuristic Q-functions*: QMDP, QPOMDP, and QBG [6].

Part of the challenge will be determine which of those planning algorithms are best suited to the Multi-Agent Challenge.

[1] http://staff.science.uva.nl/ faolieho/index.php?fuseaction=software.madp

5 Conclusion

The Multi-Agent Challenge is strongly aligned with the theoretical approach of the UvA Rescue Team, so the probability is high that our team will fully concentrate on this challenge for the competition in João Pessoa, Brazil.

References

1. H. Akin, N. Ito, A. Kleiner, J. Pellenz and A. Visser, "RoboCup Rescue Robot and Simulation Leagues", *AI Magazine*, volume 34(1):pp. 78–86, 2013.
2. A. Kleiner, A. Farinelli, S. Ramchurn, B. Shi, F. Maffioletti and R. Reffato, "RMASBench: benchmarking dynamic multi-agent coordination in urban search and rescue", in "Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems", pp. 1195–1196, International Foundation for Autonomous Agents and Multiagent Systems, 2013.
3. F. Maffioletti and R. Reffato, "Guidelines to integrate a DCOP algorithm for RMAS Benchmark", .
4. R. Nair, M. Tambe, M. Yokoo, D. Pynadath and S. Marsella, "Taming decentralized POMDPs: Towards efficient policy computation for multiagent settings", in "Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)", pp. 705–711, 2003.
5. F. A. Oliehoek, J. F. Kooi and N. Vlassis, "The Cross-Entropy Method for Policy Search in Decentralized POMDPs", *Informatica*, volume 32:pp. 341–357, 2008.
6. F. A. Oliehoek, M. T. J. Spaan and N. Vlassis, "Optimal and Approximate Q-value Functions for Decentralized POMDPs", volume 32:pp. 289–353, 2008.
7. M. Pfingsthorn, B. Slamet, A. Visser and N. Vlassis, "UvA Rescue Team 2006; RoboCup Rescue - Simulation League", in "Proc. CD of the 10th RoboCup International Symposium", 2006.
8. S. B. M. Post, M. L. Fassaert and A. Visser, "The high-level communication model for multiagent coordination in the RoboCupRescue Simulator", in "7th RoboCup International Symposium", (edited by D. Polani, B. Browning, A. Bonarini and K. Yoshida), *Lecture Notes on Artificial Intelligence*, volume 3020, pp. 503–509, Springer-Verlag, 2004.
9. M. T. J. Spaan and F. A. Oliehoek, "The MultiAgent Decision Process toolbox: software for decision-theoretic planning in multiagent systems", in "AAMAS Workshop on Multi-agent Sequential Decision Making in Uncertain Domains", pp. 107–121, 2008.
10. M. T. J. Spaan and N. Vlassis, "Perseus: Randomized Point-based Value Iteration for POMDPs", *Journal of Artificial Intelligence Research*, volume 24:pp. 195–220, 2005.
11. M. E. Taylor, M. Jain, P. Tandon, M. Yokoo and M. Tambe, "Distributed Online Multi-Agent Optimization Under Uncertainty: Balancing Exploration and Exploitation", *Advances in Complex Systems*, volume 14(03):pp. 471–528, 2011.
12. M. Traichioiu and A. Visser, "UvA Rescue - Team Description Paper - Agent competition - Rescue Simulation League - RoboCup 2014 - João Pessoa - Brazil", February 2014.
13. A. Visser, N. Dijkshoorn, S. van Noort, O. Zwennes, M. de Waard, S. Katt and R. Rozeboom, "UvA Rescue - Team Description Paper - RoboCup 2012", June 2012.