ResQ Freiburg: Deliberative Limitation Of Damage

Michael Brenner, Alexander Kleiner, Mathias Exner, Markus Degen, Manuel Metzger, Timo Nüssle, Ingo Thon

Institut für Informatik, Universität Freiburg, 79110 Freiburg, Germany

brenner, kleiner, exner, degen, mmetzger, nuessle, thon@informatik.uni-freiburg.de

Abstract. RoboCupRescue is a large-scale multiagent simulation of urban disasters. In order to save lives and minimize damage, rescue teams must effectively cooperate despite sensing and communication limitations. To accomplish this, ResQ Freiburg uses a twofold strategy: Platoon agents show reactive, yet cooperative behavior, but this can be overridden by deliberative high-level decisions of the center agents. To make these decisions, a new multiagent planning method is used in abstract search spaces generated by agent-specific clustering methods.

1 Introduction

The history of search and rescue services has repeatedly shown the difficulty of the coordination of rescue teams in disastrous situations. The RoboCupRescue domain is an attempt to simulate such situations, particularly the aspects of a dynamically changing environment, e.g. the distribution of fire, bandwidth limitation for communication and the high demand for collaborative sensing and acting in a large scale environment.

The task is to successfully coordinate teams of agents for the sake of the limitation of damage to people and buildings. There are two classes of agents: platoon and center agents. Platoon agents are small teams that operate in the field, for example a fire brigade, and center agents are stations that may instruct the teams in the field with new tasks. Groups of platoon agents take one out of three roles, namely fire brigade, ambulance team and police force. To each kind one center agent is associated.

The domain is a real multi agent scenario since most problems are not solvable by an agent on its own. Fire brigades, for example, rely on blocked roads cleared by police forces in order to reach their target. Buildings are extinguished more efficiently by more than one team, particularly if the fire spreads out in many directions. Moreover, the task is challenging due to limited communication bandwidth, the agents' limited field of view and the difficulty to predict how disasters evolve over time, e.g. how the structure of the city influences the burning speed and direction of the fire.

We implemented an agent development kit (ADK) that provides methods for communication, cooperative world modeling and path planning. Communication is carried out efficiently by compressing and routing messages via the center agents to the platoon agents which improve on their part their local world model with this information. Path planning is done by an extended A* method that takes uncertainty into account. Based on the ADK, specific skills for each agent type are implemented, such as *extinguishFire* or *clearBlockedRoad*. The decision about the execution of skills is decomposed on the one hand into a reactive part by the platoon agents and on the other hand into a deliberative part by the center agents. Platoon agents select the skill to be executed by means of hierarchical reinforcement learning [4] with respect to the current state of the environment. Central agents decide skill execution in the long term by allocating groups of platoon agents to particular tasks. Their decision making is based on a module for state prediction and abstraction that generates the input for a novel multi-agent planner [1].

The remainder of this paper is structured as follows. In Section 2 the inter-agent communication is described in more detail. Section 3 sketches the abilities of platoon and Section 4 the abilites of the center agents, respectively. In Section 5 we summarize our approach and give an outlook to the further development of our team.

2 Communication

For being in one line with real world disasters, the communication of platoon agents in the simulation is limited to the reception and transmission of four messages during each round respectively. Therefore it is necessary to restrict and compress the amount of broadcasted messages as far as possible. We utilized two methods in order to comply with this requirement. Firstly, each agent memorizes messages perceived during cycles in the past. Secondly, the central agents act as message distributors between central and platoon agents.

The first method is realized by an object pool attached to each agent. At start up, the object pool is initialized with static objects of the map, such as houses and streets. During runtime, perceived messages due to dynamic changes, such as fire or blockades, are stored with a timestamp into the data structure and transmitted to the center agent if they were not stored in the pool beforehand. Center agents collect messages transmitted by the platoon agents during each cycle. They filter and merge all collected information into one message that then is transmitted to the other center agents and the platoon agents of the same type.

Furthermore, there exists a special kind of messages for task allocation. Central agents may send an open request for a specific problem to a group of agents. The fire station, for example, might transmit a list of n buildings to a group of m platoon agents situated close to the particular site. Subsequently, the group sends a positive or negative acknowledgment, indicating whether it is possible to perform the task. Central agents may also update an open request e.g. extend the set of buildings that have to be extinguished.

3 Platoon Agents

All platoon agents have the following capabilities:

- World Modeling: Acting efficiently in a chaotic environment like a disaster area is mainly a matter of gathering information and keeping it accurate in spite of high dynamics. The world model each agent maintains is updated not only by its sensory information, but also by the “observation report” compiled by the center agent. Thus agents can share an nearly consistent global world model.
- Path Planning: A road segment is believed to be either blocked, unblocked or in an unknown state. Depending on what an agent wants to accomplish different versions of an A^*-based path-planning algorithm allow for *safe* planning (use only definitely unblocked roads) or *exploratory* planning (use unblocked or unknown roads).

- Exploration: Agents not having tasks to execute will explore the city to gather information and/or find new opportunities to act. The map is dynamically divided into areas (quadrants) such that agents can disperse and explore larger areas of the city.
- Skill selection: All capabilities (skills) of agents are described in a standardized fashion using pre- and postconditions. The precondition is a test that must be fulfilled for the skill to be applicable, the postcondition is another test describing when the skill has successfully been executed. Both conditions correspond to logical descriptions used by the planning component when selecting action sequences for agent teams [1] (cf. Sec. 5). When there is no command from the planning agents or the environment such that a plan does no longer make sense, agents will not stay idle but chose a skill to execute autonomously. The function for selecting among the applicable skills is learned by a hierarchical reinforcement learning method [4].
- Acting individually, in team, or on command: all agents have action selection schemes that allow them to chose their actions individually based on their current view of the world. However, those decisions can be overruled when external coordination is necessary: platoons can receive *commands* from their center agent and change their behavior accordingly. For example, fire agents can be commanded to act as a group and extinguish the boundaries of a fire.

3.1 Fire brigades

Fire can only be prevented from spreading by coordinated group action: neither a single fire brigade can extinguish any but the smallest fire nor can many agents crowding in the same small area. Therefore, teams are built, either assigned by the fire station (cf. Sec. 4) or emerging from an implicit, independent coordination process in which each agent chooses actions based on the perception of other agents nearby. In a simple distributed role assignment process every agent in the team assumes a specific position and role: some fire brigades detect fire boundaries and disperse to control its spreading while others try to extinguish sites where civilians are particularly endangered.

3.2 Ambulance Teams

Ambulance teams search for injured civilians inside buildings. Besides their own perception they are informed by messages from their center agent about observations from other agents in the field. By this, there exist at each cycle n targets to be rescued by m agents. The ambulance teams prefer to dig out civilians with critical injuries, however, not in case a rescue seems to be hopeless. This is decided by comparing the current state, e.g. *hitpoints* , of the target and the time a rescue would take, e.g. the degree of *buriedness*. In order to prevent multiple agents from rescuing the same target, the center agent also acts as an arbiter between the platoon agents.

3.3 Police Forces

Police forces must clear blocked roads to enable other agents to reach their destinations. Therefore, a heuristic ranking of roads to be cleared should be based on their possible use for other agents. The ResQ police forces employ several heuristics to rank roads, some computed offline, i.e. when the city map is first made public, and some computed during the rescue mission. One offline method predicts importance of road segments by calculating shortest paths between all pairs of nodes on the map. This can be done efficiently with the Floyd-Warshall algorithm [2] (see also [3]). Another offline heuristic increases priorities of roads leading to refuges (important both for civilians seeking shelter and, newly, fire brigades refilling their tanks). Online heuristics include increasing evaluation for roads leading to fire sources and decreasing factors when other police agents are closer to a blocked road.

4 Center Agents

Center Agents are responsible for message routing (see Section 2) and long-term task allocation. Successful task allocation depends strongly on a good prediction of the disaster's extension and the teams' capability to limit the damage. The spreading of fire, for example, depends on the fuel density, e.g. density of houses, the number of floors of each house or the material houses are made of. In order to predict the fire's extent in the future, one has to take these factors into account. Another important problem is given by blocked roads due to collapsing buildings. Here the inter-connectivity of regions on the map is of particular interest.

We currently examine clustering methods that learn online the neighborhood relations of buildings and roads in order to tackle these problems. Locations on the map are clustered to regions that are reachable within one cycle, whereas buildings are clustered in terms of their inter-inflammability relation.

Based on these predictive models, we utilize a multi-agent planner that decides the simultaneous deployment of teams in order to maximize the limitation of damage in a minimal amount of time [1].

5 Summary and outlook

In this paper we introduced the cooperative and deliberative capabilities of the ResQ Freiburg team, particularly cooperative sensing by message routing and task allocation in the long term. At the current stage, deliberation is still very limited in our system. In the future, we want to extend both planning and learning capabilities within the context of hierarchical data structures.

References

1. Michael Brenner. Multiagent planning with partially ordered temporal plans. Submitted to IJCAI'03.
2. T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*. MIT Press, 1992.
3. J. Habibi, M. Ahmadi, A. Nouri, M. Sayyadian, and M. Mohammadi. Utilizing different multiagent methods in robocuprescue simulation. Team Description, 2002.
4. A. Kleiner, M. Dietl, and B. Nebel. Towards a life-long learning soccer agent. In *Proc. Int. RoboCup Symposium '02*. Fukuoka, Japan, 2002.