Utilizing Different Multiagent Methods in RoboCupRescue Simulation

Jafar Habibi Mazda Ahmadi Ali Nouri
Mayssam Sayyadian Mayssam Mohammadi

Department of Computer Engineering,
Sharif University of Technology, Tehran, Iran
arian@ce.sharif.edu http://ce.sharif.edu/robocup/arian

Abstract. This paper describes the design and implementation of a MultiAgent System in the field of RoboCupRescue Simulation. Some traditional and new algorithms regarding autonomous agents and multi agent systems are used in the system. Some aspects of these algorithms will be briefly described.

1 Introduction

Arian team is the result of a research group in Robocup field in Sharif University of Technology(SUT). We chose RobocupRescue as a research field for the sake of humanity. We hope someday our research can help someone live longer.

RoboCupRescue Simulation with heterogeneous agents with different abilities and of course responsibilities with a limited communication, is an excellent framework for multiagent planning, communication techniques, coalition formation and task allocation.

In RoboCupRescue there are three types of agents and three head quarters. In this paper the internal algorithms for these three types of agents and methods for communication between the different types of agents is described. These algorithms contain a reinforcement learning method for police forces, a task allocation method based on combinatorial auctioneer problem for firebrigades, and another reinforcement learning method for ambulance and a message evaluation method used for communication between agents.

This paper is organized as follows: In section 2, the overview of the system is presented. In section 3 the algorithms of fire brigade agent will be explored. In section 4 police force algorithms will be described. In section 5 the techniques used in ambulance will be presented. In sections 6 and 7 the methods for path planning and communication between heterogeneous agents are presented and finally section 8 will conclude the paper.

2 overview of RobocupRescue Simulation

RoboCup Rescue Simulation is one of the competitions in Robocup. The main aim of RoboCup Rescue Simulation is simulating a disaster situation in a city.

There is a kernel simulating the city and some simulators simulating the disaster conditions [2],[7]. The parts that we have developed are the agents. The agents are:

- Fire Brigades
- Ambulance Teams
- Police Forces
- Head Quarters

The main goal of the agents is to rescue more civilians. Although ambulances are responsible for rescuing civilians, but polices will clear the roads so that ambulances and fire brigades can move in the city. Fire brigades have to extinguish fires to reduce the amount of damage (the less fire, the more civilians).

3 Fire Brigade

Our firebrigade agents first consider sites of fire and try to stop spreading fire by extinguishing the boundary of the site.

It worths noting that buildings can be extinguished much faster if multiple firebrigades are assigned to a fiery building. So, our firebrigades tend to form coalitions to extinguish fiery buildings. To find an optimal formation of coalitions we consider firebrigades as resources which shall be allocated to some buildings. Some works have been done before on Collaborative Planning with similar task allocation problem [4] and some beneficial algorithms have been developed [3]. This way the problem can be viewed from the auctioneer point of view [6] it can be said that buildings bid for agents.

To find a decentralized approximate solution for the winner determination problem, we use some sort of algorithm to find an approximation of the result and take advantage of local improvements using steepest hill climbing.

When evaluating the value of each bid, the factors are as follows:

- Each building has a priority based on its strategic position and situation in the city and the fire site boundary. This priority will be calculated regarding to number of buildings which can be ignited in the following cycles by that building and some other factors.
- Time cycles it takes a coalition of firebrigades to extinguish the building.

4 Police Force

The task of polices is to clear roads. Trying to clear more roads is not the optimum action. Polices have to select the most important roads. Importance of the road is how many times other agents will pass through this road in the following cycles. although the exact value of this measure is not computable before the simulation. So we compute the probable value, for this we computed shortest path between all the city nodes, and those city nodes which apeared in more paths have more priority. for this reason we used floyd algorithm [Cormen].

There is main decision for the polices to make. they are:

- stay around the current place and clear around.
- Go to help another agent who have asked for help
- Go to clear a critical region (An example of a critical region is around fires)

Polices have to decide what to do. For this reason a reinforcement learning method (QLearning) [5] is used. Actions are as mentioned. Rewards are given after clearing a road proportional to the roads' priority. State of the system is constituted from the following parameteres:

- Number of other agents near us.
- Normalized distance of the places we can go (if any)
- Number of clearances in the last 10 cycles. (It's a measure of usefullness in the region)
- Number of cycles I have been near the current place.

We have learned the system about 150 times (450'000 Cycles). It seems that the functionality of the police has improved with this learning method.

5 Ambulance Team

Ambulance Team agents rescue injured civilians. They obtain information of civilians by means of communication, and gathering audiotorial and visual information. In order to determine whether to go to rescue a civilian or move around to find an injured civilian, we've used a QLearning method.

6 Path Planning

As it was proposed, the second phase of each agent action is to find a path toward it's goal. In the rescue simulation area, where about 70% of the roads are cut off by the earthquake and mostly, the agents have the information of only one road, use of a combinatorial algorithm of path planning with a high rate of communication is inevitable. We have adjusted our agents so that they have some kind of correlation in using the paths in the city (i.e. blind communication). To achieve this, we have devided the roads into two categories containing super roads and normal roads. The idea is to minimize a path cost by passing it through the super roads as much as we can. At the same time, the police forces will find these super roads and clear them in early times so other agents have a spanning tree of these roads in the city. Beside this blind communication, normal communication is also used as a supplementary method to gain extra information about the roads.

7 Communicatin Method

In our communication system every agent who asks for help, sets a priority in that message. This priority indicates the importance of the situation and how accurate the information is.

One of the problems in communication is to derive a value for the messages. we derived an equation for evaluating the message. which is:

$$V_i(t) = \frac{f(t)*P(i)}{succ(t-1)}$$

In the above equation $V_i(t)$ indicates the value of i'th message in each time period and ofcourse each agent has it's own queue of messages and has it's own $V_i(t)$. $f(t)$ is a coefficient which is related to time and so makes $V_i(t)$ related to time. $f(t)$ is what we are aiming to learn and has a constant value at initial state. by learning $f(t)$ we can say that $V_i(t)$ is learned. $P(i)$ indicates the priority that the message has. $succ(t-1)$ is the measure of the success of the agent in last time period. If this value of $f(t)$ is less than a threshold we will accept the message and do what the message has told us to do, if not we will ignore that message.

we use an incremental function approximation method for learning the value of $f(t)$. in our function approximation method we use an initial value for $f(t)$ and in each cycle we update the $f(t)$ function. The way we update this function has an explicit relation with the agent abilities and the environment of the problem. In the following subsections we introduce our functions such as $succ(t)$, $P(i)$ and the way we update $f(t)$.

8 Conclusion and Future Work

In this paper some different algorithms for different agents in RoboCupRescue Simulation have been presented breifly. Testing other reinforcement algorithms such as Monte Carlo (instead of QLearning) with other state parameters is among our future works. We Also consider testing a coalition formation method [6] for firebrigades. The communicatoin system needs some refinement in the case of determining priority on the side of the agent who wants to send a message. We are also planning to use a QLearning algorithm on this part.

References

[1] Thomes H Cormen, Charles E, leiserson, Ronald L. Rivest, Introduction to Algorithems, MIT Press, 1992.

[2] The RoboCupRescue Technical committee, RoboCupRescue Simulator Manual.

[3] Thomas Sandholm, et al. A Fast Optimal Algorithm for Combinotorial Auctions

[4] Luke Hunsberger, Barbara J. Grosz, A Combinatorial Auction for Collaborative Planing, IC-MAS2000

[5] Richard S. Suttun, Andrew G. Barto Reinforcement Learning: An Introduction, MIT Press, Cambridge, MA, 1998

[6] O. Shehory and S. Kraus, Coalition formation among autonomous agents: Strategies and complexity, Lecture Notes in Artificial Intelligence no. 957, From Reaction to Cognition, C. Castelfranchi and J. P. Muller (Eds.), pages 57-72, 1995.

[7] The Robocup Federation, Robocup Regulations and Rules, http://www.robocup.org.