# RoboCup Rescue Simulation: Methodologies Tools and Evaluation for Practical Applications

A. Farinelli, G. Grisetti, L. Iocchi, S. Lo Cascio, D. Nardi

Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Salaria 113, 00198 Roma, Italy
E-mail : `<lastname>@dis.uniroma1.it`

**Abstract.** The activities of search and rescue of victims in large-scale disasters are not only highly relevant social problems, but pose several challenges from a scientific standpoint. In this context, the RoboCup-Rescue project focused on the problems of bringing aids immediately after a large disaster, and aims at creating system based on AI and Robotics technologies, where heterogeneous agents (software, robots, human beings) interact in a cooperative manner.
In this paper we present the achievements of a research project, based on the RoboCup Rescue simulator, carried out in Italy in collaboration with the Italian Fire Department. The overall project goal is to devise tools to allow monitoring and supporting decisions which are needed in a real-time rescue operation in a large scale disaster, and to provide a methodology for evaluation of multi-agent system which considers not only the efficiency of a system, but also its robustness when conditions in the environment change, as well as other features, such as the ability to acquire a precise and coherent representation of the disaster scenario.

## 1 Introduction

Search and rescue of victims in large-scale disasters is a very relevant social problem, and pose several challenges from a scientific standpoint. When earthquakes, eruptions or floods happen, a considerable organizational capability to aid the disaster victims as fast as possible is required. This task is rather difficult since often different secondary disasters (e.g. fires, damages in the transportation and communication systems) connected with the main one, occur, which make the correct execution of a rescue plan a priori decided impossible. Moreover, one of the critical issues for a proper response to emergencies is to have timely and reliable information on the disaster scenario, thus enabling for a more effective use of the resources available for the rescue operations [3].

In the recent past significant research initiatives have been undertaken in Japan [5], in the USA [3], and also in Italy [4], that specifically focus on the problem of developing software tools to support the management of this kind of emergency and, more specifically, to design a support system for search and rescue operations in large-scale disasters. These tools are intended both for on-line support during the actual operations as well as for previsional analysis and training.

In this paper we present the achievements of a research project [4], based on the RoboCup Rescue simulator, carried out in Italy in collaboration with the Italian Fire Department and the University of Geneva and Politecnico di Milano. The overall goal of the project was to devise tools to allow monitoring and supporting decisions which are needed in a real-time rescue operation in a large scale disaster.

The activities needed to reach the project's goals were: (i) the creation of a simulation environment, (ii) the design of a framework for Cognitive Agent Development and (iii) a definition of a evaluation methodology. Since in this paper we want to focus more on the evaluation aspects for a MAS in the rescue domain than on the agent development, the agent development tools are just sketched, in order to give a basic for understanding the experimental settings. The first two issues are described in the next section, the evaluation methodology is outlined in section 3.

## 2 Project Description

In this section we provide a general overview of the project "Real-time planning and monitoring for search and rescue operations in large-scale disasters"[1], that has been developed in collaboration with other Italian Universities (University of Genova and Politecnico di Milano) and the VVF (Italian Fire Department) [4].

The goal of this project has been to develop a prototype tool, based on the RoboCup Rescue simulator, to allow monitoring and supporting decisions which are needed in a real-time rescue operation in a large scale disaster, by integrating competences and tools already available to the VVF and using as a case-study event the Marche and Umbria earthquake in Fall 1997.

The development of a system with the desired characteristics requires to integrate in an effective way, the three following main components: (1) modeling of events related, in a direct way or not, to the disaster; (2) acquiring and integrating data coming from different heterogeneous sources; (3) modeling/monitoring/planning the resources used in the intervention. The simulator developed in the framework of RoboCup-Rescue considers simultaneously all of the three elements above-mentioned, offering an environment for experimentation, which provides completely innovative characteristics compared to those of the existing applications related to this field. For these reasons the RoboCup Rescue simulator has been chosen as a basic tool in the project development for monitoring and planning rescue operations.

In order to ground our project in a real scenario we have chosen to access the data of the Umbria and Marche earthquake (1997), and in particular of the city of Foligno, that is one of the most important cities in that region. Foligno is located in a flat region of eastern Umbria. Its urban structure is characterized by a medieval center surrounded by more recent suburbs; in particular we focused our attention on an area of about 1 $km^2$ in the city center. In the area under consideration there are no high-rise buildings; most recent structures are mid-rise, in the four- to nine-story range. All large, multi-story buildings were constructed of reinforced concrete. Oldest buildings were mainly constructed of rubble-work, whereas only few structures are steel frame buildings or wood

[1] Funded under the program Agenzia 2000 of the Italian Consiglio Nazionale delle Ricerche.

buildings. There are no industrial structures; most buildings are housing estates having variously-shaped plants. The road network is quite irregular, with not very large roads and narrow alleys (see Figure 1).

Moreover, by considering and analyzing the structures and the strategies currently adopted by the Fire Department (VVF), we have acquired not only a significant body of expertise on the rescue operations, but also the opportunity to set up a prototype experimental setting to show the results of the project.

In order to reach the project goal, two main operations have been performed:

– the set up of an initial environment situation, simulating a disaster in Foligno,
– the definition of multi-agent systems acting in the simulated world performing rescue operations.

In the following we shortly describe two tools that address the above issues, while in Section 3 we focus on the experiments that has been preformed with the systems created on those tools. Such tools have to be considered in this context, mainly for the support given in the rapid development of the MAS systems on which the experiments have been run, and for a better understanding of the settings.

*GIS Editor* . For running the RoboCup Rescue simulator on a new scenario, a map of the chosen simulation environment in a format suitable for the simulator is needed, moreover some other parameters, like the fire ignition points, or the earthquake magnitudo level, has to be specified. In order to match the above requirements we realized a graphical tool, for generating RoboCup Rescue configuration files: GIS Editor, a screenshot of which is shown in Figure 1.

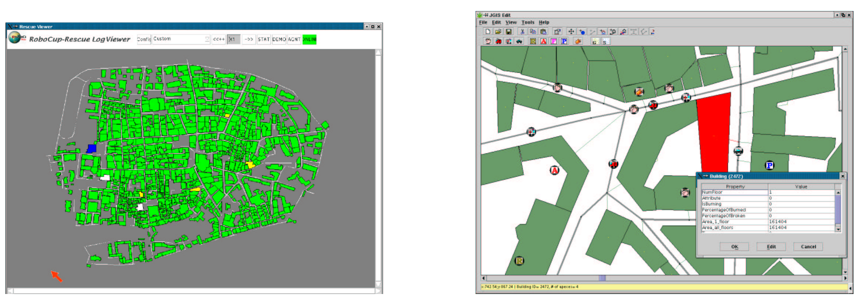


**Fig. 1.** The map of Foligno(right). A GIS editor screenshot(left)

*Cognitive Agent Development Kit.* In order to reach good performances in the post-earthquake disaster situation agents needs to exhibit both planning and cooperation capabilities, since the abilities of a single individual agent are often not enough for fighting an expanding disaster. Another item to be considered while developing a team of rescue agents is the need of integrating partial and noisy information coming from the agents, in order to assess a global situation, on which to perform the resource allocation. The Cognitive Agent Development Kit (CADK) is a tool that allows for the rapid development of a team of agents that perform rescue operations in the simulated world. Moreover the CADK provides some debugging facilities for run-time inspecting the internal agent state and the communication state. We used the CADK for experimenting different combination of methods for approaching the above problems in a structured way.

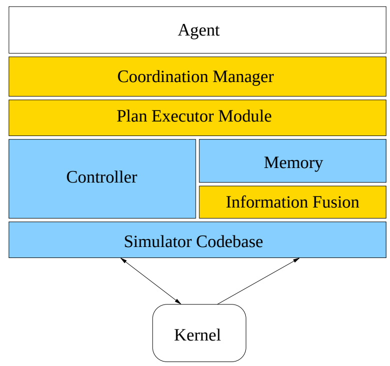


**Fig. 2.** The Cognitive Agent Development Kit architecture

The architecture of an agent built on the CADK is sketched in Figure 2. The CADK itself is built on the RoboCup Rescue Agent Development Kit [1], by adding coordination, planning and information integration layers.

## 3 Experimental evaluation of multi-agent systems in the RoboCup Rescue domain

Evaluation of multi-agent systems in the RoboCup Rescue domain is important not only within the RoboCup Rescue simulation competitions, but also for evaluating actual plans to be used during rescue emergencies. Evaluation of MAS in the RoboCup Rescue domain is currently carried out within international contests, by rating each competing rescue team with a score representing the result of its activity in a simulated scenario. The one with the highest measured score is the contest winner.

In the real world, however, events not always develop in a known and predictable way, since unexpected charges in the operative conditions and failures can occur at every time. The evaluation rule used in RoboCup Rescue simulation competitions is applied in a standard fixed situation and does not take into account the ability of a multi-agent system to work under troublesome situations and its ability to adapt to non-standard operative conditions.

The evaluation method proposed here is based on [2] and allows the analysis of a rescue team in a more realistic way, by analyzing the performance of a multi-agent system in terms of efficiency under normal conditions, as well as in terms of reliability and robustness under changing working conditions.

### 3.1 Experimental setting

To acquire a measure of the reliability and the robustness of a MAS, a series of simulations have been executed under changing operative conditions. These tests give a measure of the system adaptability to unexpected situations. The changing parameters that we have considered are: (i)perception radius, (ii)number of agents and (iii) errors in the communication system. Each parameter characterizes a particular series of simulations, referred respectively as the visibility test, the disabled agents test and the noisy communications test.

**The visibility test.** In outdoor environments, visibility conditions are extremely variable. Rescue operation can be needed every time of the day, also in the night. Thus, it is necessary to probe the activity of a system also in these situations. The visibility test is performed by executing five simulations, each with decreasing perception radius, modeling activities under different visibility conditions (i.e. twilight, night time, fog). In this test the changing conditions are on the perception range of each agent, that is 30 meters under normal conditions, and we have performed experiments for the same multi-agent system also using 20, 10, 5 and 3 meters of perception range.

**The disabled agents test.** In a real emergency situation, it can happen that an agent suddenly become not operational for some reason (for example a mechanical failure of its vehicle or its equipment); this test analyzes the reactions of a system against new operative conditions in which some of the operative agents are disabled.

The disabled agents test is composed of five simulations: in the normal conditions all the agents are active, for the other simulations one, two, three, and four of the *best* agents for each force are disabled. The choice of the best agents to disable is based on the number of tasks performed: for each force, the agent that has completed more tasks in shorter time will be disabled.

**The noisy communication test.** Agent cooperation in the rescue domain is mainly attained by radio communications among coordination centers and between a coordination center and the operative agents. In real conditions communication transmissions are not free from network failures, or human misunderstandings. This test verifies the robustness of an analyzed multi-agent system by introducing errors in the communication channel, thus preventing messages to reach their destination. The noisy communication test is composed of five phases: under normal conditions there are no errors in the communication channel, while in the other simulations 1/10, 1/3, 1/2, and 9/10 of the sent messages are lost.

### 3.2 Performance measures

The performance of a rescue multi-agent system is measured in terms of efficiency and reliability. The **efficiency** is directly evaluated by the formula used in RoboCup-Rescue tournaments, which is:

$$V = (P + S/S_0) * \sqrt{B/B_0}$$

where $P$ is the number of living agents, $S$ is the remaining hit points (health level) of all agents, $S_0$ is the total hit points of all agents at initial, $B$ is the area of houses that are not burnt and $B_0$ is the total area at initial; the higher the value of $V$ for a rescue system, the better the results of the rescue operation.

The **reliability** describes how much system efficiency is affected by the worsening of operative conditions, and how much it depends on the values $V$ assumed in the simulation sequence of a single test. It is evaluated with the linear regression slope formula:

$$LRS = \frac{\sum_{i=0}^{N-1}(x_i - x_m) * (y_i - y_m)}{\sum_{i=0}^{N-1}(x_i - x_m)^2}$$

where $(x_i, y_i)$ are the coordinates of a point in a Cartesian system, $(x_m, y_m)$ the average values of these coordinates, $N$ the number of points considered.

Since each point of the graph represents the value of $V$ obtained in the i-th phase, this formula is applied for evaluating the reliability of a system, by using $x_i = i$ and $y_i = V(i)$. For example, if in the visibility test $V$ assumes values described in the following table:

| Simulation | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Visibility | 30m | 20m | 10m | 5m | 3m |
| $V$ | 23.3 | 12.6 | 17.1 | 10.7 | 11.0 |

the efficiency is 23.3, while the reliability, calculated by the LRS formula, is -2.65 . It is reasonable to have a negative value of the LRS, since usually the effectiveness of the agents decreases with the worsening of the operative conditions. A little absolute value of the LRS means a good degree of reliability of the system to adverse situations.

### 3.3 Performance comparison

The values presented in the previous section are of little significance if not compared to the measures obtained from simulations of other rescue systems. The performance comparison allows to establish the effectiveness of a new technique over the previous ones or over the state-of-the-art.

In the following section it is shown as an example the performance evaluation executed on four different rescue-systems created with the ADK tool. The analyzed MAS are distinguished for the information integration and resource allocation techniques employed, as shown in the following table:

| | Simple fusion integration | No fusion integration |
|---|---|---|
| Static resource allocation | MAS1 | MAS2 |
| Dynamic resource allocation | MAS3 | MAS4 |

For these four multi-agent systems, the controlled experimentations are executed, providing the results shown in the following tables:

| Visibility test | Effic. | Rel. |
|---|---|---|
| MAS1 | 23.3 | -2.65 |
| MAS2 | 18.8 | -2.06 |
| MAS3 | 18.8 | -1.99 |
| MAS4 | 11.6 | -1.67 |

| Disabled agents | Effic. | Rel. |
|---|---|---|
| MAS1 | 23.3 | -6.18 |
| MAS2 | 18.8 | -5.23 |
| MAS3 | 18.8 | -5.09 |
| MAS4 | 11.6 | -4.88 |

| Noisy communic. | Effic. | Rel. |
|---|---|---|
| MAS1 | 23.3 | -3.18 |
| MAS2 | 18.8 | -3.95 |
| MAS3 | 18.8 | -4.04 |
| MAS4 | 11.6 | -2.79 |

In each test there is a rescue system which gets the best value about efficiency and another one which is better in reliability. Rarely in these tests the same rescue system is the best for the two measures, since usually more sophisticated techniques means less robustness to nonstandard operative conditions.

To get some more meanings from the previous results, a graphical representation is presented, sorting the results of each test both by Efficiency and by Reliability, as shown in the following picture.

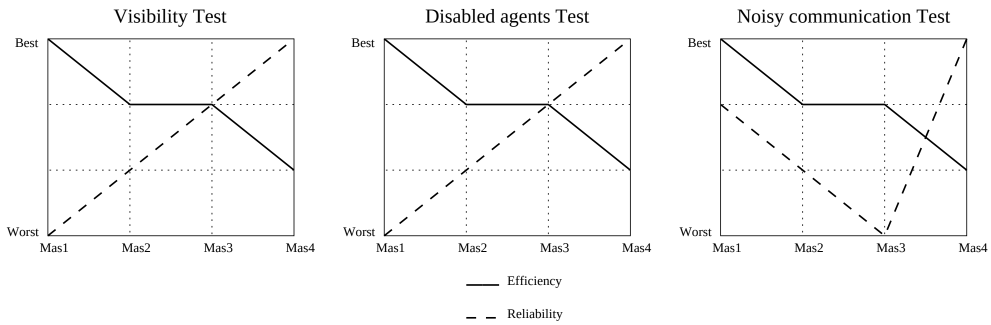


**Fig. 3.** Performance comparison

At last, it is not easy to identify which system has the best overall performance. In the first test, the visibility one, MAS1 is the best one in terms of efficiency, but it gets the worst rating about reliability. MAS2 and MAS3 have the same efficiency value, and are jointly ranked in the second place. MAS4, which is the worst system in terms of efficiency, it's the best one with respect to reliability. Looking at the graph, it can also be seen that MAS3 offers the best compromise between efficiency and reliability, since it is second in both of the two measures. In the noisy communication test, MAS1, which has the best efficiency value, is also a good system in terms of reliability, ranking in the second place; in this case, it seems to be better than the other ones.

This example shows that the choice of the best system is hard to cast in absolute terms. Depending on the application of the evaluated systems, it should be selected the system which offers the best score with respect to efficiency, reliability, or to a (weighted) combination of the two, but this is still an open problem.

Currently, in RoboCup Rescue tournaments, competing teams are evaluated for their activities in standard situations, thus measuring only their efficiency. It should be pointed out that rescue teams in real life barely operate in normal conditions, and official tournaments should also evaluate the ability of a rescue system to face unattended situations.

## 4 Conclusions

We have presented the current development of the project "Real-time planning and monitoring for search and rescue operations in large-scale disasters". The aim of this research has been to develop a tool to support search and rescue operations in large scale disasters. In particular, we built our system upon the RoboCup Rescue Simulator, that provides an environment for experimentation of multi-agent technology in the framework of the RoboCup initiative. We have addressed a specific application domain: the disaster scenario recorded after the earthquake of Umbria and Marche. The availability of the RoboCup simulator has been extremely valuable for the development of the present project, providing an experimental setting that can be effectively used for developing a prototype implementation.

The main results of the project has been:

1. a set of tools for improving the development of rescue systems based on the RoboCup Rescue simulator, namely: the GIS Editor and the Cognitive Agent Development Kit.
2. the definition of an evaluation system for MAS in order to analyze different rescue strategies

The research developed in this project has provided a significant use of agent technology in the design of tools supporting the acquisition of information as well as the planning of activities when there is the need to act immediately with partial information about the situation, as in a typical emergency scenario.

Evaluating the activity of a MAS is an hard task since it involves the analysis of several parameters, and the system evolves in unpredictable ways; moreover the rescue operations has to be performed in situations in which the agent adaptability to different environment settings is required. The proposed evaluation system allows for an easy comparison of different MAS, providing an easy understandable graphical representation of the performances, and considers the robustness a key factor as well as the agent efficiency under normal conditions for the overall performance measure.

While the prototype developed has been designed for demonstration and not intended for actual operation, it shows the potential benefits of an integrated approach to the simulation and monitoring of a real search and rescue scenario. While it is premature to consider the effectiveness of the tool in the management of operation, both the analysis of past scenarios as well as the training of personnel seem to be already suitable for application.

## References

1. Micheal Bowling. Robocup rescue: Agent developement kit. *available at official RoboCup Rescue Web Site,*, 2000.
2. G. A. Kaminka, I. Frank, K. Arai, and K. Tanaka-Ishii. Performance competitions as research infrastructure: Large scale comparative studies of multi-agent teams. *Journal of Autonomous Agents and Multi-Agent Systems*, 2002. To appear.
3. J. Llinas. Information fusion for natural and man-made disasters. In *Proc. of 5th International Conference on Information Fusion*, 2002.
4. D. Nardi, A. Biagetti, G. Colombo, L. Iocchi, and R. Zaccaria. Real-time planning and monitoring for search and rescue operations in large-scale disasters. In *Technical Report University "La Sapienza" Rome*, 2002. `http://www.dis.uniroma1.it/~rescue/`.
5. S. Tadokoro and et al. The robocup rescue project: a multiagent approach to the disaster mitigation problem. *IEEE International Conference on Robotics and Automation (ICRA00), San Francisco*, 2000.