USARSim: A RoboCup Virtual Urban Search and Rescue Competition

Stephen Balakirsky[a], Chris Scrapper[a], Stefano Carpin[b], and Mike Lewis[c]

[a]National Institute of Standards and Technology, Gaithersburg, MD
[b]University of California Merced, Merced, CA
[c]University of Pittsburg, Pittsburg, PA

ABSTRACT

Research efforts in Urban Search And Rescue (USAR) robotics have grown substantially in recent years. A virtual USAR robotic competition was established in 2006 under the RoboCup umbrella to foster collaboration amongst institutions and to provide benchmark test environments for system evaluation. In this paper we describe the software simulation framework that is used in this competition and the rules and performance metrics used to determine the league's winner. The framework allows for the realistic modeling of robots, sensors, and actuators, as well as complex, unstructured, dynamic environments. Multiple heterogeneous agents can be concurrently placed in the simulation environment thus allowing for team or group evaluations.

Keywords: USAR, Multi-agent, Robotics, Competition, RoboCup, USARSim, Performance Metrics

1. INTRODUCTION

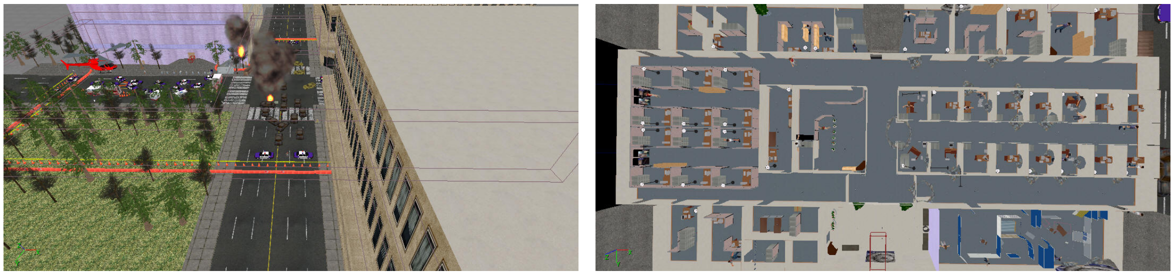
Urban Search and Rescue (USAR) presents an extremely challenging arena for robotics. The structural diversity of a disaster area presents obstacles that make the use of a robot with limited capabilities problematic. For example, there are places only reachable by climbing robots, spots only accessible through small openings, long passageways that may need to be traversed, and regions only observable by aerial vehicles.

In general the problem is not solvable by a single agent, and a heterogeneous team that dynamically combines individual capabilities in order to solve the task is needed.[1] Multi-robot teams not only offer the possibility to meet these challenges, they also exhibit increased robustness due to redundancy,[2] and superior performance thanks to parallel task execution.[3] However, research into the use of multi-robot teams exposes researchers to many unwanted side effects. It may be prohibitively expensive to obtain the necessary robots and sensors, maintenance becomes a substantial burden, a large enough environment to employ such a team may become an issue, and standardized techniques for team evaluation are often lacking. The virtual USAR robotic competition was established to provide the infrastructure necessary to address these issues.

This virtual competition falls under the umbrella of the RoboCup Rescue competitions. The RoboCup Rescue competitions provide a benchmark for evaluating robot platforms for their usability in disaster mitigation and are experiencing ever increasing popularity.[4] Roughly speaking, the *league vision* can be paraphrased as the ability to deploy teams of robots that explore a devastated area and locate victims. Farsighted goals include the capability to identity hazards, provide structural support and more. RoboCup Rescue is structured in two leagues, the Rescue Robot League and the Rescue Simulation League. Whereas the Rescue Robot League fosters the development of high-mobility platforms with adequate sensing capabilities, e.g. to identify human bodies under harsh conditions, the Rescue Simulation League promotes research in planning, learning, and information exchange in an inherently distributed rescue effort.[4] The Rescue Simulation League contains three competitions; the Virtual Robot Competition, the Agent Competition, and the Infrastructure Competition. The competition discussed in this paper is the Virtual Robots competition.[5] The Virtual Robots competition simulates, compared to the Rescue Agents competition, small teams of agents* with realistic capabilities operating on a city block-sized scenario.

Further author information: (Send correspondence to stephen.balakirsky@nist.gov)

*The maximal team-size during the competition in 2006 was 12 virtual robots.



(a) (b)

Figure 1. Overview of the simulated world used for the 2006 RoboCup Virtual Robot competition. The world consisted of outdoor (a) and indoor (b) areas.

The Virtual Robot competition, first held during the RoboCup competitions in 2006, provides a realistic simulation environment for simulating conditions after a real disaster,[6] such as an earthquake, a major fire, or a car wreck on a highway. Robots are simulated on the sensor and actuator level, making a transparent migration of code between real robots and their simulated counterparts possible. The simulation environment, described later on in this paper, allows evaluation of the performance of large robot teams and their interactions. For example, whereas in the real robot competition there are usually only one or two robots deployed, in the Virtual Robot competition teams deployed up to twelve. Furthermore, the simulator provides accurate ground truth data allowing an objective evaluation of the robots' capabilities in terms of localization, exploration and navigation, e.g. avoidance of bumping.

In this paper we provide a detailed description of the Virtual Robot competition. An important goal for this competition is to promote research about quantitative evaluation of intelligent teams performance. For this reason we also describe the efforts aimed at the development of such metrics. Specifically, during the first edition of the competition, tools to measure the effectiveness in mapping and exploration were developed and evaluated on the field by the competing teams. We speculate that this sort of research, currently not receiving the deserved degree of attention from the scientific community, is better addressed in a simulation environment. Massive data collection to compute parameters with statistical meaning are better collected in a simulated environment. The competition held in 2006 represents the first attempt in this direction, and we report about the lessons learned.

The winning teams from the 2006 RoboCup Rescue Virtual Competition were:

- **First Place** *Rescue Robots Freiburg*, University of Freiburg, Germany
- **Second Place** *Virtual IUB*, International University Bremen, Germany
- **Third Place** *UvA*, University of Amsterdam, The Netherlands
- **Best Mapping** *UvA*, University of Amsterdam, The Netherlands
- **Best Human-Computer Interface** *Steel*, University of Pittsburgh, USA

2. PERFORMANCE METRICS

In designing metrics for the Virtual Robot competition, the objectives of developing relevant metrics for the urban search and rescue community, providing metrics that could be applied to both virtual and real robots with little modification, and assuring that the metrics did not place an undo burden on competitors were all balanced. In order to maintain relevance, metrics from the Rescue Robot League that emphasized the competency of the robots in tasks that would be necessary at a disaster site were used as a starting point. These tasks included locating victims, providing information about the victims that had been located, and developing a comprehensive map of the explored environment. The environment (shown in Figure 1) for the competition consisted of an

(a) (b)

Figure 2. Maze environment used for real/virtual development and testing. (a) is the real world maze and (b) is its virtual representation.

arena that was over 11,000 m^2 in size and contained a multi-story office building (with office space and cubicals), outside streets, a park, and several accident scenarios. Participants needed to deploy teams of agents to fully explore the environment.

While this was a fictitious world, researchers have modeled their real-world lab environments to assist in real/virtual testing and development. One such environment is shown in Figure 2. Figure 2.a shows an actual outdoor maze environment for robotic testing. Figure 2.b shows the simulated version of this environment. During testing, real robots cooperate with virtual robots to perform missions and behaviors. For example, a real robot may be exploring the maze in cooperation with a virtual robot. The robots share map information and even see each other in their own respective representations of the real or virtual worlds.

The primary goal of the competition was to locate victims in the environment and determine their health status. However, what does it meant to "locate" a victim? How does one autonomously obtain health status? Several possible interpretations exist ranging from simply requiring a robot to be in proximity of a victim (e.g. drive by the victim) to requiring the robot to employ sensor processing to recognize that a victim is located near-by (e.g. recognize a human form in a camera image) and then examine that victim for visually apparent injuries. While recognizing a human from a camera image is the solution most readily portable to a real hardware, it places an undo burden on both the competitors and the evaluation team. For the competitors, a robust image processing system would need to be developed that could recognize occluded human forms. No matter how exceptional the mapping and exploration features of a team were, failing to produce the image processing module would result in a losing effort. In addition, the evaluation team would need to develop an entire family of simulated human forms so that teams could not "cheat" by simply template matching on a small non-diverse set of victims. It was decided that robots should be required to be "aware" of the presence of a victim, but that requiring every team to have expertise in image processing was against the philosophy of lowering entry barriers. Therefore, a new type of sensor: a victim sensor, was introduced. To allow for the metrics to be portable to real hardware, this new sensor would need to be based on existing technology.

The victim sensor was based on Radio Frequency Identification Tag (RFID) technology. False alarm tags were scattered strategically in the environment, and each victim contained an embedded tag. At long range (10 m), a signal from the tag was readable when the tag was in the field of view (FOV) of the sensor. At closer range (6 m), the sensor would report that a victim or false alarm was present. At even closer range (5 m) the ID of the victim would be reported. Finally, at the closest range (2 m), the status of the victim (e.g. injured, conscious, bleeding, etc.) was available. Points were subtracted for reporting false alarms, and were awarded for various degrees of information collected from the victims. Bonus points were awarded for including an image of the victim with the report.

As the robots were exploring the environment, their poses (on 1 s intervals) and any collisions between the robots and victims were automatically logged. The pose information was fed into a program that automatically computed the amount of area that was covered by the robotic teams. This figure was normalized against the expected explored area for the particular run, and points were awarded accordingly. The collision information was used as an indication of suboptimal navigation strategies that should be penalized. Another parameter that was used to determine the overall score was the number of human operators that were needed to control the robots. The idea was borrowed from the Rescue Robot league with the intent of promoting the deployment of fully autonomous robot teams, or the development of sophisticated human-robot interfaces that allow a single operator to control many agents.

The final area that was judged during the competition was map quality. The map quality score was based on several components.

- Metric quality – The metric quality of a map was scored automatically by examining the reported locations of “scoring tags”. Scoring tags were single shot RFID tags (they could only be read once). A requirement of the competition was for the teams to report the global coordinates of these tags at the conclusion of each run. The automatic scoring program then analyzed the deviation of the perceived locations from the actual locations.
- Multi-vehicle fusion – Teams were only permitted to turn in a single map file. Those teams that included the output from multiple robots in that single map were awarded bonus points.
- Attribution – One of the reasons to generate a map is to convey information. This information is often represented as attributes on the map. Points were awarded for including information on the location, name, and status of victims, the location of obstacles, the paths that the individual robots took, and the location of RFID scoring tags.
- Grouping – A higher order mapping task is to recognize that discrete elements of a map constitute larger features. For example the fact that a set of walls makes up a room, or a particular set of obstacles is really a car. Bonus points were awarded for annotating such groups on the map.
- Accuracy – An inaccurate map may make a first responder’s job harder instead of easier. Points were assessed based on how accurately features and attributes were displayed on the map.
- Skeleton quality – A map skeleton reduces a complex map into a set of connected locations. For example, when representing a hallway with numerous doorways, a skeleton may have a line for the hallway and symbols along that line that represent the doors. A map may be inaccurate in terms of metric measurements (a hallway may be shown to be 20 m long instead of 15 m long), but may still present an accurate skeleton (there are three doors before the room with the victim). The category allowed the judges to award points based on how accurately a map skeleton was represented.
- Utility – One of the main objectives of providing a map was to create the ability for a first responder to utilize the map to determine which areas had been cleared, where hazards may be located, and where victims were trapped. Points were granted by the judges that reflected their feelings on this measure.

$$S = \frac{V_{ID} * 10 + V_{ST} * 10 + V_{LO} * 10 + t * M + E * 50 - C * 5 + B}{(1+N)^2} \tag{1}$$

The above mentioned elements were numerically combined according to Equation 1. The meaning of the variables is discussed below. This equation represents a schema that took into account merit factors that concerned (1) victims discovery, (2) mapping, and (3) exploration. The exact point calculations for each factor are presented below.

1. 10 points were awarded for each reported victim ID (V_{ID}). An additional 10 points were granted if the victim’s status (V_{ST}) was also provided. Properly localizing the victim in the map was rewarded with an additional 10 points (V_{LO}). At the referee’s discretion, up to 20 bonus points were granted for additional information produced (B). For example, some teams managed to not only identify victims, but to also provide pictures taken with the robot’s cameras. For this additional information teams were awarded with 15 bonus points.

(a) (b)

Figure 3. Representative snapshot of a USARSim indoor (a) and outdoor (b) scene.

2. Maps were awarded up to 50 points based on their quality (M), as previously described. The obtained score was then scaled by a factor ranging between 0 and 1 (t) that measured the map's metric accuracy. This accuracy was determined through the use of the RFID scoring tags.

3. Up to 50 points were available to reward exploration efforts (E). Using the logged position of every robot, the total amount of explored square meters (m^2) was determined and related to the desired amount of explored area. This desired amount was determined by the referees and was based on the competition environment. For example, in a run where 100 m^2 were required to be explored, a team exploring 50 m^2 would receive 25 points, while a team exploring 250 m^2 would receive 50 points, i.e. performances above the required value were leveled off.

On the penalization side, 5 points were deducted for each collision between a robot and a victim (C). Finally, the overall score was divided by $(1 + N)^2$, where N was the number of operators involved. So, completely autonomous teams, i.e. $N = 0$, incurred no scaling, while teams with a single operator had their score divided by 4. No team used more than one operator.

It should be noted that except for the map quality, all of the above components were automatically computed from the information logged during the competition. Therefore subjective opinions during the scoring stage were reduced to a minimum. In an ideal scenario, the scoring step would be completely automatic as is currently the case for the RoboCup Rescue Agents competition. In addition to assigning points to determine the overall best systems, the judges assigned winning teams in the special categories of map creation and human-machine interface. The map creation award was presented to the team that consistently scored the highest in the map quality assessment while the human-machine interface award recognized the team with the most innovative robot control console. These performance metrics were successfully applied to judging the 2006 Virtual Robot competition at RoboCup 2006.

3. VIRTUAL ROBOT COMPETITION

The RoboCup Rescue Virtual Robot Competition is the third competition running under the RoboCup Rescue Simulation League umbrella. It utilizes the USARSim framework to provide a development, testing, and competition environment that is based on a realistic depiction of a disaster scenario. It has been previously stated[7,8] that the Virtual Robot competition should serve the following goals:

- Provide a meeting point between the different research communities involved in the RoboCup Rescue Agents competition and the RoboCup Rescue Robot League. The two communities are attacking the same problem from opposite ends of the scale spectrum (city blocks vs. a small rubble area) and are currently far apart in techniques and concerns. The Virtual Robot competition offers close connections to the Rescue Robot League, as well as challenging scenarios for multi-agent research. The scenarios for the 2006 competition were chosen to highlight these connections. They were based on an outdoor accident scene, and an indoor fire/explosion at an office building. These scenarios included real-world challenges such as curbs, uneven terrain, multi-level terrain (i.e. the void space under a car), maze-like areas, stairs, tight spaces, and smoke. An exact copy of one of the RoboCup Rescue Robot League arenas was also included in the office space, and elements of other arenas were scattered throughout the environment. The area was far too large to be explored by a single agent in the time permitted (20 minutes) and thus the use of multi-agent teams was beneficial. Accommodations were provided in the worlds to assist less capable (in terms of mobility) robotic systems. For example, wheel-chair ramps were provided that allowed for alternative access around stairs. Snap shots of small sections of these environments may be seen in Figure 3.

- Lower entry barriers for newcomers. The development of a complete system performing search and rescue tasks can be overwhelming. The possibility to test and develop control systems using platforms and modules developed by others makes the startup phase easier. With this goal in mind, the open source strategy already embraced in the other competitions is fully supported in the Virtual Robot competition. Software from this year's top teams has already been posted on the web.
- Let people concentrate on what they can do better. Strictly connected to the former point, the free sharing of virtual robots, sensors, and control software allows people to focus on certain aspects of the problem (victim detection, cooperation, mapping, etc), without the need to acquire expensive resources or develop complete systems from scratch. In order to help people determine if they really can "do better," performance metrics were applied to the competing systems.

4. USARSIM FRAMEWORK

The current version of USARSim[9] is based on the UnrealEngine2 game engine that was released by Epic Games as part of Unreal Tournament 2004.[†] This engine may be inexpensively obtained by purchasing the Unreal Tournament 2004 game. The engine handles most of the basic mechanics of simulation and includes modules for handling input, output (3D rendering, 2D drawing, and sound), networking, physics and dynamics. Multiplayer games use a client-server architecture in which the server maintains the reference state of the simulation while multiple clients perform the complex graphics computations needed to display their individual views. USARSim uses this feature to provide controllable camera views and the ability to control multiple robots. In addition to the simulation, a sophisticated graphical development environment and a variety of specialized tools are provided with the purchase of Unreal Tournament.

The USARSim framework[‡] builds on this game engine and consists of:

- standards that dictate how agent/game engine interaction is to occur,
- modifications to the game engine that permit this interaction,
- an Application Programmers Interface (API) that defines how to utilize these modifications to control an embodied agent in the environment,
- 3-D immersive test environments.

When an agent is instantiated through USARSim, three basic classes of objects are created that provide for the complete control of the agent. These include robots, sensors, and mission packages and are defined as part of the API to USARSim. For each class of objects there are class-conditional messages that enable a user to

[†]Certain commercial software and tools are identified in this paper in order to explain our research. Such identification does not imply recommendation or endorsement by the authors or their institutions, nor does it imply that the software and tools tools identified are necessarily the best available for the purpose.

[‡]The USARSim framework can be downloaded from `http://sourceforge.net/projects/usarsim`.

query the component's geography and configuration, send commands, and receive status and data. Permissible calls into the game engine and complete details on the API may be found in the USARSim Reference Manual.[10]

It is envisioned that researchers will utilize this framework to perfect algorithms in the areas of:

- Autonomous multi-robot control,
- Human, multi-robot interfaces,
- True 3D mapping and exploration of environments by multi-robot teams,
- Development of novel mobility modes for obstacle traversal,
- Practice and development for real robots that will compete in the RoboCup Rescue Robot league.

Moreover, it is foreseeable that USARSim will also be valuable in robotics education contexts.[11]

5. RESULTS

		RRFreiburg	**GROK**	**IUB**	**SPQR**	**STEEL**	**UvA**
	# operators	**0**	1	**0**	1	1	**0**
Preliminary 1	Victim points	70	20	20	10	**100**	0
	Map points	-	16	19	-	22	**24**
	Exploration points	**50**	8	18	49	**50**	12
	Total points	120	44	57	59	**172**	36
	Operator corrected points	**120**	11	57	15	43	36
Preliminary 2	Victim points	20	10	20	55	**70**	5
	Map points	-	15	10	19	**21**	8
	Exploration points	**50**	15	26	48	44	26
	Total points	70	40	56	122	**135**	39
	Operator corrected points	**70**	10	56	30	34	39
Preliminary 3	Victim points	75	75	50	10	**160**	30
	Map points	-	**31**	24	7	27	-
	Exploration points	**50**	15	41	11	31	30
	Total points	125	120	115	28	**218**	60
	Operator corrected points	**125**	30	115	7	54	60
Semifinal 1	Victim points	170	140	260	203	**285**	150
	Map points	-	23	29	15	**47**	42
	Exploration points	**50**	5	38	16	22	44
	Total points	220	168	327	235	**354**	235
	Operator corrected points	**220**	42	327	59	89	235
Semifinal 2	Victim points	380	125	205	315	**410**	40
	Map points	12	39	26	36	**45**	9
	Exploration points	**50**	21	35	31	35	**50**
	Total points	442	185	267	382	**489**	99
	Operator corrected points	**442**	46	266	95	122	99
Final 1	Victim points	**245**	-	145	-	-	-
	Map points	37	-	**48**	-	-	-
	Exploration points	**50**	-	**50**	-	-	-
	Total points	**332**	-	243	-	-	-
Final 2	Victim points	**270**	-	135	-	-	-
	Map points	46	-	**47**	-	-	-
	Exploration points	**50**	-	34	-	-	-
	Total points	**366**	-	216	-	-	-

Table 1. Final results from RoboCup '06

This section overviews the results of the RoboCup Rescue Virtual competition 2006. The competition was held in three preliminary rounds, two semifinals and two finals. The scoring follows the metrics described in Section 2 and thus evaluates team performance at *victim discovery, mapping*, and *exploration*. Table 1 reports the overall scores for each team during the entire competition. In the following, the performance measure respectively for victim discovery, mapping and exploration are detailed.

Providing information about the victims is the primary task of the robot team. Therefore, the victim score is not bound to a maximum, but grows linearly with the number of victims found. Bonus points were awarded both for locating victims and providing extra information. Table 2 summarizes the victim information reported

		RRFreiburg	GROK	IUB	SPQR	STEEL	UvA
Preliminary 1	reported victims	**5**	2	1	1	2	0
	status bonus	**20**	0	10	0	**20**	0
	victim bonus (picture, etc)	-	-	-	-	-	-
Preliminary 2	reported victims	2	1	1	**3**	2	1
	status bonus	20	0	10	**30**	20	0
	victim bonus (picture, etc)	-	-	15	-	**30**	-
Preliminary 3	reported victims	**5**	3	3	1	**5**	3
	status bonus	**50**	0	0	0	**50**	0
	victim bonus (picture, etc)	-	45	30	-	**60**	-
Semifinal 1	reported victims	**8**	4	7	6	7	6
	status bonus	**70**	20	30	50	60	30
	victim bonus (picture, etc)	-	60	**105**	45	90	-
	localization bonus	**80**	20	70	48	70	60
Semifinal 2	reported victims	**16**	4	5	9	14	2
	status bonus	80	0	40	60	**140**	0
	victim bonus (picture, etc)	-	45	**75**	**75**	0	-
	localization bonus	**160**	40	50	90	140	20
Final 1	reported victims	**10**	-	3	-	-	-
	status bonus	**60**	-	30	-	-	-
	victim bonus (picture, etc)	-	-	**60**	-	-	-
	localization bonus	**100**	-	30	-	-	-
Final 2	reported victims	**10**	-	5	-	-	-
	status bonus	**80**	-	10	-	-	-
	victim bonus (picture, etc)	-	-	**30**	-	-	-
	localization bonus	**100**	-	50	-	-	-

Table 2. Victim results from RoboCup '06

by each team. In particular, the table shows the number of victims identified, the bonus for reporting the status (i.e 10 points for each victim status), and extra information such as pictures (up to 20 points for each report) and accurate localization of the victims (up to 20 points). It is clear from Table 2 that RRFreiburg found at every round the largest number of victims. Still, STEEL received the largest amount of bonus points due to additional information they provided about the victims. The human operators could use the camera of the robot to get detailed information about the victims which was not available to the automated teams.

Exploration was evaluated based on the totally explored area of each team. These values were automatically computed from the logs of the server hosting the simulator. Table 5 gives an overview on the number of deployed robots, and area explored by each team. Furthermore, the average area explored by a single robot of each team is shown (*Area/#robots*). RRFreiburg demonstrated that it is possible to quickly explore a large area with a team or robots. On average, their robots drove 5 times as fast as the robots of the other teams. Detailed mapping of the disaster area is not possible at those speeds, but reliable navigation could be performed.

From the results it is clear that no team dominated in all the tasks. In particular, the winning team Rescue Robots Freiburg was the only team which addressed cooperative-exploration explicitly, whereas most of the "best" teams adopted a selfish approach. Their RFID-based exploration, allowed to quick and efficient exploration of large environments at the cost of abstracting the environment representation to a topological map. On the other hand, teams such as UvA, focused on providing high quality merged maps using state of the art techniques which resulted in a third place and best mapping award. UvA thus had to spend more time assessing map quality, rather than exploring, which penalized them in the final scoring both for exploration and number of found victims. VirtualIUB, which reached the second place, mainly focused on building a complete system. They integrated several existing robotic approaches outperforming teams that underestimated some of the robotic challenges. The migration of well known robotic approaches towards USARSim, in general, shows that Virtual Robot Competition is a meaningful test-bed. Finally, STEEL which won the best Human-Robot Interface award, was the only team in the "best" four that was not fully autonomous. It is clear from the

		RRFreiburg	GROK	IUB	SPQR	STEEL	UvA
Preliminary 1	# Robots	12	1	6	4	6	1
	Area [m2]	**902**	31	70	197	353	46
	Area/#robots	**75,17**	31	11,67	49,25	58,83	46
	Speed [m/s]	**0,96**	0,19	0,20	0,46	0,50	0,47
Preliminary 2	# Robots	12	1	4	4	6	8
	Area [m2]	**550**	61	105	191	174	104
	Area/#robots	45,83	**61**	26,25	47,75	29	13
	Speed [m/s]	**1,09**	0,55	0,31	0,44	0,29	0,19
Preliminary 3	# Robots	10	1	5	7	6	7
	Area [m2]	**310**	59	164	44	124	120
	Area/#robots	31	**59**	32,8	6,29	20,67	17,14
	Speed [m/s]	**0,67**	0,48	0,42	0,06	0,32	0,35
Semifinal 1	# Robots	8	1	6	4	6	6
	Area [m2]	**579**	27	227	96	134	262
	Area/#robots	**72,38**	27	37,83	24	22,33	43,67
	Speed [m/s]	**2,62**	0,33	0,35	0,30	0,27	0,51
Semifinal 2	# Robots	8	1	6	5	6	7
	Area [m2]	**1276**	82	139	123	139	286
	Area/#robots	**159,5**	82	23,17	24,6	23,17	40,86
	Speed [m/s]	**2,61**	0,66	0,21	0,21	0,38	0,48
Final 1	# Robots	8	-	8	-	-	-
	Area [m2]	**1203**	-	210	-	-	-
	Area/#robots	**150,38**	-	26,25	-	-	-
	Speed [m/s]	**2,67**	-	0,24	-	-	-
Final 2	# Robots	8	-	6	-	-	-
	Area [m2]	**350**	-	136	-	-	-
	Area/#robots	**43,75**	-	22,67	-	-	-
	Speed [m/s]	**1,82**	-	0,36	-	-	-

Table 3. Exploration Results from RoboCup '06

results that their performance was impressive although they were strongly penalized by the scoring, for having an operator. In fact, their approach allowed their only operator to successfully coordinate up to seven robots.

6. CONCLUSIONS

This paper presented the outcome of the first Virtual Robot competition at RoboCup 2006. It discussed preliminary efforts of defining performance metrics in terms of mapping, localization, and exploration, for robot teams operating in the rescue scenario. The wide spectrum of solutions, ranging from autonomy to teleoperation, and from detailed grid-based map representations to abstract topological map representations, shows clearly that there exists a real need for evaluating aspects of team collaboration in this domain, i.e. to identify methods for combining the individual team strengths in an ideal way.

The encouraging outcome of the competition gives rise to multiple research and development directions that will be pursued in the next contests. First of all, despite the already proved close correspondence between results obtained with simulated robots inside USARSim and with the corresponding physical robots, the necessity to obtain more validation data cannot be overemphasized. In particular, a standardized procedure to validate simulated robots is needed to avoid the addition of poorly validated models to the simulation environment.

While the competition provided an overall winner that can claim "best in search and rescue", it proved to be difficult to discern the exact reasons behind the win. Was it better mapping, better mobility, or better strategy? To address this, next year's competition will include a round of elemental tests that will provide insite into individual skills and competencies. The tests will include a mapping task, a mobility challenge, and a victim finding challenge. These individual tests will allow us to declare a best in class in each of the skills necessary for urban search and rescue.

Other additions to the competition are being made due to the necessity of closing the loop between simulation and reality. Therefore, the use of a software module to mimic wireless communication between agents is being introduced. When physical robots are fielded in a disaster scenario, their efforts are useless if collected information is not made available to human operators. A preliminary extension of USARSim implementing this idea has already been developed and is included in the official distribution. The new communication model accounts for the signal strength of wireless transmitters and spacial structure of buildings, thus denying communication over long distances and communication through impenetrable walls. Hence, the new model requires teams to implement coordination strategies for sharing, and more importantly, reporting back their information to a central command post.

In addition to sharing communications, steps are being taken to encourage heterogeneous teams in this year's competition (all teams last year employed homogeneous teams of robots). To accomplish this a small payload flying platform and a slow, large payload, high-mobility platform are introduced. It is hoped that teams will construct heterogeneous groups of robots and apply each of their strong points to addressing the various challenges.

Another important milestone towards heterogeneous robot teams, which is currently under discussion within all Rescue leagues, is to introduce a standardized map representation. A common format could serve not only for comparisons between different approaches, but also as useful tool to exchange spatial information between heterogeneous agents, as well as to compare more rigorously results coming from simulation and real world systems. Furthermore, it makes possible to provide floor plans of building structures to rescue teams in advance, as is common practice during real rescue missions. Here, the long term goal is to introduce communication standards in the Virtual competition and to transfer them to the Rescue Robot league, after they have been extensively tested. Therefore, in future competitions all maps will be submitted to the judges in GEOTIF (a pixel format) or MIF (a vector format). This will allow the judges to more readily overlay the generated maps on ground truth to evaluate their quality, as well as to provide floor plans in advance.

On a more general perspective, it appears also that a systematic approach to team work evaluation is still lacking in the multi-robot research community. The use of a high fidelity simulator like USARSim opens the doors for more thorough investigation in this direction without the need of cumbersome or expensive experimental facilities. The RoboCup event, on the other side, offers the perfect venue to contrast different approaches in a competitive scenario and verify if the developed metrics capture strengths and weaknesses of the developed solutions.

REFERENCES

1. R. Murphy, J. Casper, M. Micire, and J. Hyams, "Mixed-initiative control of multiple heterogeneous robots for usar," Tech. Rep. CRASAR-TR2000-11, Center for Robot Assisted Search & Rescue, University of South Florida, Tampa, FL, 2000.
2. M. Dias, R. Zlot, M. Zinck, and A. Stentz, "Robust multirobot coordination in dynamic environments," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, **4**, pp. 3435–3442, 2004.
3. T. Arai, E. Pagello, and L. Parker, "Guest editorial advances in multirobot systems," *IEEE Transactions on Robotics and Automation* **18**(5), pp. 655–661, 2002.
4. H. Kitano and S. Tadokoro, "Robocup rescue: A grand challenge for multiagent and intelligent systems.," *AI Magazine* **22**(1), pp. 39–52, 2001.
5. S. Carpin, M. Lewis, J. Wang, S. Balakirsky, and C. Scrapper, "Bridging the gap between simulation and reality in urban search and rescue," in *RoboCup 2006: Robot Soccer World Cup X*, *LNAI*, Springer, 2006.
6. J. Wang, M. Lewis, S. Hughes, M. Koes, and S. Carpin, "Validating usarsim for use in hri research," in *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, pp. 457–461, 2005.
7. S. Carpin, M. Lewis, J. Wang, and S. Balakirsky, "Bridging the gap between simulation and reality in urban search and rescue," in *Robocup 2006: Robot Soccer World Cup X*, *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 2007. (to appear).

8. S. Carpin, J. Wang, M. Lewis, A. Birk, and A. Jacoff, "High fidelity tools for rescue robotics: results and perspectives," in *Robocup 2005: Robot Soccer World Cup IX*, *Lecture Notes in Computer Science* **4020**, pp. 301–311, Springer Berlin / Heidelberg, June 2006.
9. S. Balakirsky, C. Scrapper, S. Carpin, and M. Lewis, "Usarsim: Providing a framework for multi-robot performance evaluation," in *Proc. of the Performance Metrics for Intelligent Systems (PerMIS) Workshop*, pp. 98–102, August 2006.
10. J. Wang and S. Balakirsky, "Usarsim v 3.0." `http://sourceforge.net/projects/usarsim`, 2007.
11. S. Carpin, M. Lewis, J. Wang, S. Balarkirsky, and C. Scrapper, "Usarsim: a robot simulator for research and education," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2007. (to appear).