A Reference Test Course for Urban Search and Rescue Robots

Adam Jacoff, Elena Messina, John Evans

National Institute of Standards and Technology, Intelligent Systems Division
100 Bureau Drive, MS 8230
Gaithersburg, MD 20899-8230
adam.jacoff@nist.gov

Abstract

One approach to measuring the performance of intelligent systems is to develop standardized or reproducible tests. These tests may be in a simulated environment or in a physical test course. The National Institute of Standards and Technology has developed a test course for evaluating the performance of mobile autonomous robots operating in an urban search and rescue mission. The test course is designed to simulate a collapsed building structure at various levels of fidelity. The course will be used in robotic competitions, such as the American Association for Artificial Intelligence (AAAI) Mobile Robot Competition and the RoboCup Rescue. Designed to be highly reconfigurable and to accommodate a variety of sensing and navigation capabilities, this course may serve as a prototype for further development of performance testing environments. The design of the test course brings to light several challenges in evaluating performance of intelligent systems, such as the distinction between "mind" and "body" and the accommodation of high-level interactions between the robot and humans. We discuss the design criteria for the test course and the evaluation methods that will be used.

Introduction

The Intelligent Systems Division of the National Institute of Standards and Technology (NIST) is researching how to measure the performance of intelligent systems. One approach being investigated is the use of test courses for evaluating autonomous mobile robots operating in an urban search and rescue scenario. Urban search and rescue (USAR) is an excellent candidate for deploying robots, since it is an extremely hazardous task. USAR refers to rescue activities in collapsed building or man-made structures after a catastrophic event, such as an earthquake or a bombing. Japan has an initiative, based on the RoboCup robots, that focuses on multi-agent approaches to the simulation and management of major urban disasters (Kitano et al. 1999). The real-world utility and manifold complexities inherent in this domain make it attractive as a "challenge" problem for the mobile autonomous robots community (Blitch 1996, Casper and Murphy 2000).

The type of environments that a rescuer has to confront with collapsed buildings is shown in Figure 1. There is totally unstructured rubble, which may be unstable and contain many hazards. Victims' locations and conditions must be established quickly. Every passing minute reduces the chances of saving a victim.

This type of environment stresses the mobility, sensing, and planning capabilities of autonomous systems. The robots must be able to crawl over rubble, through very narrow openings, climb stairs or ramps, and be aware of the possibility of collapses of building sections. The sensors are confronted with a dense, variable, and very rich set of inputs. The robot has to ascertain how best to navigate through the area, avoiding hazards, such as unstable piles of rubble or holes, yet maximizing the coverage. The robot also has to be able to detect victims and ideally, determine their condition and location. The robot has to make careful decisions, planning its path and strategy, and taking into account the time constraints.

A near-term measure of success for robots in a search and rescue mission would be to scout a structure, map its significant openings, obstacles, and hazards, and locate victims. The robots would communicate with victims, leaving them with an emergency kit that contains a radio, water, and other supplies, and transmit a map, including victim locations and conditions, to human supervisors. Humans would then plan the best means of rescuing the victims, given the augmented situational awareness.

Search and rescue missions are not amenable to teleoperation due to the fact that most of the radio frequencies are reserved by emergency management agencies. Obstructions and occlusions also diminish the effectiveness of radio transmissions. Tethers are not typically practical in the cluttered environment in which these robots must operate.

Figure 1: Partial and Total Collapse of Buildings after Earthquake

USAR as a Robotic Challenge

A search and rescue mission is extremely challenging and dangerous for human experts. This is a highly unstructured and dynamic environment, where the mission is time critical. Very little a priori information about the environment or building may exist. If any exists, it will almost certainly be obsolete, due to the collapse.

Urban search and rescue is therefore attractive as a mission framework in which to measure intelligence of autonomous robots. The high degree of variability and unpredictability demand high adaptation and sophisticated decision-making skills from the robots. Robots will need to quickly and continually assess the situation, both in terms of their own mobility and of the likelihood of locating more victims. USAR missions are amenable to cooperation, which can be considered another higher-level manifestation of intelligence. We propose that any robot or team of robots that is able to successfully and efficiently carry out USAR missions would be considered intelligent by most standards.

In the following sections, we will briefly discuss how USAR missions tax specific components of an intelligent system.

Mobility

As can be seen in Figure 1, the mobility requirements for search and rescue robots are challenging. They must be able to crawl over piles of rubble, up and down stairs and steep ramps, through extremely small openings, and take advantage of pipes, tubes, and other unconventional routes. The surfaces that they must traverse may be composed of a variety of materials, including carpeting, concrete blocks, wood, and other construction material. The surfaces may also be highly unstable. The robot may destabilize the area if it is too heavy or if it bumps some of the rubble. There may be gaps, holes, sharp drop-offs, and discontinuities in the surfaces that the robot traverses.

Sensing

In order to be able to explore USAR sites and successfully navigate in this environment, the robot's sensing and perception must be highly sophisticated. Lighting will be variable and may be altogether missing. Surface geometry and materials may absorb emitted signals, such as acoustic signals, or they may reflect them. For truly robust perception, the robots should emulate human levels of vision.

The presence of victims may be manifested through a variety of signals. The stimuli that the robots have to be prepared to process include:

- Acoustic – victims may be calling out, moaning softly, knocking on walls, or otherwise generating sounds. There will be other noises in the environment due to shifting materials or coming from other USAR entities.
- Thermal – a victim will emit a thermal signature. There may be other sources of heat, such as radiators or hot water.
- Visual – a multiplicity of visual recognition capabilities, based on geometric, color, textural, and motion characteristics, will be exercised. Recognizing human characteristics, such as limbs, color of skin, clothing is important. Motion of humans, such as waving, must be detected. Confusing visual cues may come from wallpaper, upholstery or curtain material, strewn clothing, and moving objects, such as curtains blown by a breeze.

Knowledge Representation

In order to support the sophisticated planning and decision-making that is required, the robot must be able to leverage a rich knowledge base. This entails both a priori expertise or knowledge, such as how to characterize the traversability of a particular area, as well as gained information, such as a map that is built up as it explores. It must develop rich three-dimensional spatial maps that contain areas it, or other robots, have and haven't yet seen, victim and hazard locations, and potential quick exit routes. The maps from several robots may need to be shared and merged.

A variety of types of knowledge will be required in order to successfully accomplish search and rescue tasks. Higher-level knowledge, which may be symbolic, includes representations of what a "victim" is. This is a multi-facetted definition, which includes the many manifestations that imply a victim's presence.

Planning

An individual robot must be able to plan how to best cover the areas it has been assigned. The time-critical nature of its work must be taken into account in its planning. It may need to trade off between delving deeper into a structure to find more victims and finding a shortcut back to its human supervisors to report on the victims it has already found.

Autonomy

As mentioned above, it is not currently practical to assume that the robots will be in constant communication with human supervisors. Therefore, the robots must be able to operate autonomously, making and updating their plans independently. In some circumstances, there may be limited-bandwidth communications available. In this case, the robots may be able to operate under a mixed-initiative mode, where they have high-level interactions with humans. The communications should be akin to those that a human search and rescue worker may have with his or her supervisor. It definitely would not be of a teleoperative nature.

Collaboration

Search and rescue missions seem ideally suited for deploying multiple robots in order to maximize coverage. An initial strategy for splitting up the area amongst the robots may be devised. Once they start executing this plan, they will revise and adapt their trajectories based on the conditions that they encounter. Information sharing between the robots can improve their efficiency. For example, if a robot detects that a particular passageway that others may need to use is blocked, it would communicate that to its peers. The robots should therefore collaborate and cooperate as they jointly perform the mission. They may be centrally or decentrally controlled. The robots themselves may all have the same capability, or they may be heterogeneous, meaning that they have different characteristics. Heterogeneous robot teams may apply the marsupial approach, where a larger robot transports smaller ones to their work areas and performs a supervisory function.

Measuring the Performance of USAR Robots

We have described briefly the requirements for autonomous urban search and rescue robots. We will now discuss approaches to testing their capabilities in achieving USAR missions.

The approach being taken by the USAR robot competitions that use the NIST test course are based on a point system. The goal of the robots is to maximize the number of victims and hazards located, while minimizing the amount of time to do so and the disruption of the test course.

Specifically, the AAAI Mobile Robot competition (http://www.aic.nrl.navy.mil/~schultz/aaai2000) will use Olympic-style scoring. Each judge will have a certain number of points that can be awarded based on their measuring certain quantitative and qualitative metrics.

Robots receive points for:

- Number of victims located
- Number of hazards detected
- Mapping of victim and hazard locations
- Staying within time limits
- Dropping off a package to victims representing first aid, a radio, or food and water
- Quality of communications with humans
- Tolerance of communications dropout

Robots lose points for:

- Causing damage to the environment, victims, or themselves (e.g., destabilizing a structure)
- Failing to exit within time limits

In certain sections of the test course, robots are allowed to have high-level communications with humans. These communications must be made visible to the judges. Metrics for evaluating the quality of the communication include "commands" per minute and/or bandwidth used. Fewer commands per minute and less bandwidth per minute receive better scores. Tolerance of communications disruption is an important capability and will be given greater difficulty weighting. A team may request that the judges simulate communication disruptions at any point in order that the robots demonstrate how to recover. Examples of recovery would be to move to a location where there is better chance of communication, making decisions autonomously instead of consulting humans, or utilizing companion robots to relay the information to the humans.

For teams consisting of multiple robots, the advantage of cooperating or interacting robot must be demonstrated. This can be either in performing the task better, or performing the task more economically. Multi-robot teams should have a time speedup that is greater than linear, or may be able to perform the tasks with less overall power consumption or cost. The scoring will factor in the number of robots, types of robots, types or mixture of sensors, etc., in determining the performance of a team.

The RoboCup Rescue competition (http://www.robocup.org/games/36.html), sponsored by Robot World Cup Initiative, takes an evaluation benchmarking approach. Initially, there are three benchmark tasks. The current tasks are victim search, victim rescue, and a combination of victim search and rescue. Additional ones will be added as the competition and participants evolve. The RoboCup Rescue includes a simulation infrastructure in which teams can compete, as well as the use of the NIST test course.

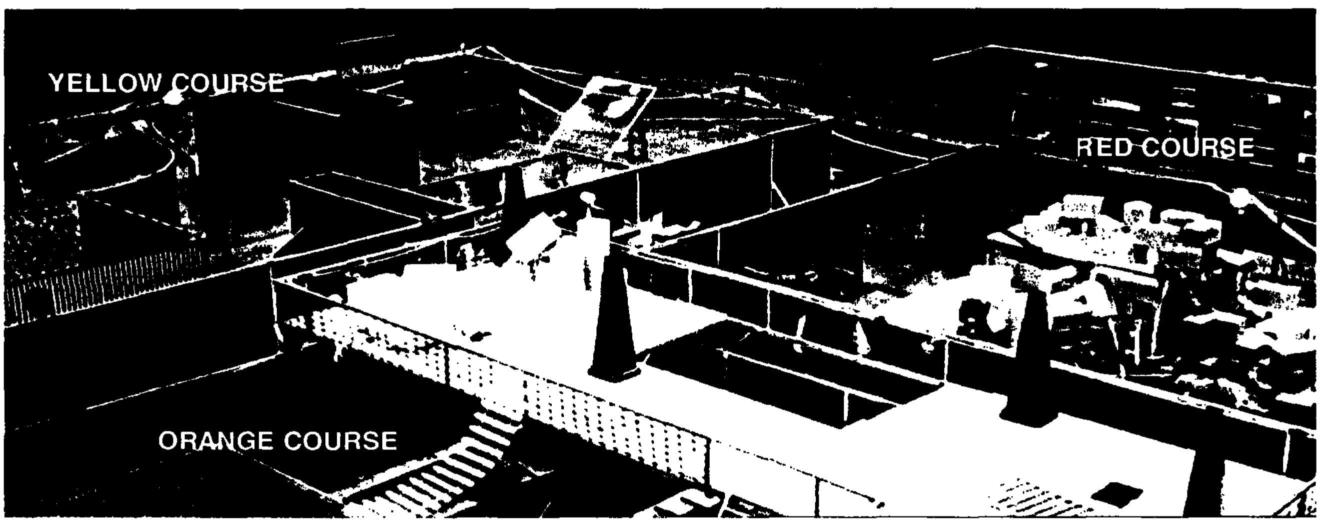

Figure 2: NIST's Reference Test Course as used for the AAAI 2000 Mobile Robot Competition

Their evaluation metrics are still under development. Examples of criteria that have been published on their web site include:

- Recovery rate, expressed as percentage of victims identified versus number under the debris.
- Accuracy rate, computed as the number of correctly identified victims divided by the total number of identified victims.
- Operational loading, which is the number of operations that a human has to perform in order to enable to robots to perform their tasks.
- If rescuing victims, the total time it takes to rescue all victims.
- Total damage caused to victims in attempting to rescue them.

The Test Course Design

The test course that NIST designed for the AAAI Mobile Robot Competition (Figures 2 and 3) was designed with three distinct areas of increasing difficulty designated the Yellow, Orange and Red Courses. Overall, the course is meant to represent several of the sensing, navigation, and mapping challenges that exist in a real USAR situation. As discussed above, these are challenges that correlate well with general characteristics desirable in mobile, autonomous robots that may operate in other types of missions. In the design of the course, trade-offs were made between realism and reproducible, or controlled, conditions. In order to be able to evaluate the performance of robots in specific skill areas, certain portions of the course may look unrealistic or too simplified. This idealization is necessary in order to abstract the essential elements being exercised, such as the ability to deal with a particular sensing challenge.

Given the controlled conditions the test course provides, it is possible to have multiple robots or teams face the identical course and have their performances compared. This should yield valuable information about what approaches to robotic sensing, planning, and world modeling work best under certain circumstances.

The course is highly modular, allowing for reconfiguration before and during a competition. Judges may swap wall panels that are highly reflective for some that are fabric-covered, for example, or victims may be relocated. This reconfigurability can serve to avoid having robot teams "game" the course, i.e., program their robots to have capabilities tailored to the course they've seen previously. The reconfigurability can serve to provide more realism as well. A route that the robot used previously may become blocked, forcing the robot to have to find an alternative way.

The three areas of the course are described below. A representative layout of the entire course is shown in Figure 3 without the added visual complexity, textures, debris and targets the robots must cope with during competition. The entire course covers roughly 400 square meters (20 m. x 20 m.). Note that the use of color in the names of the section is for labeling purposes only and does not mean that the courses are primarily colored in their namesake color.

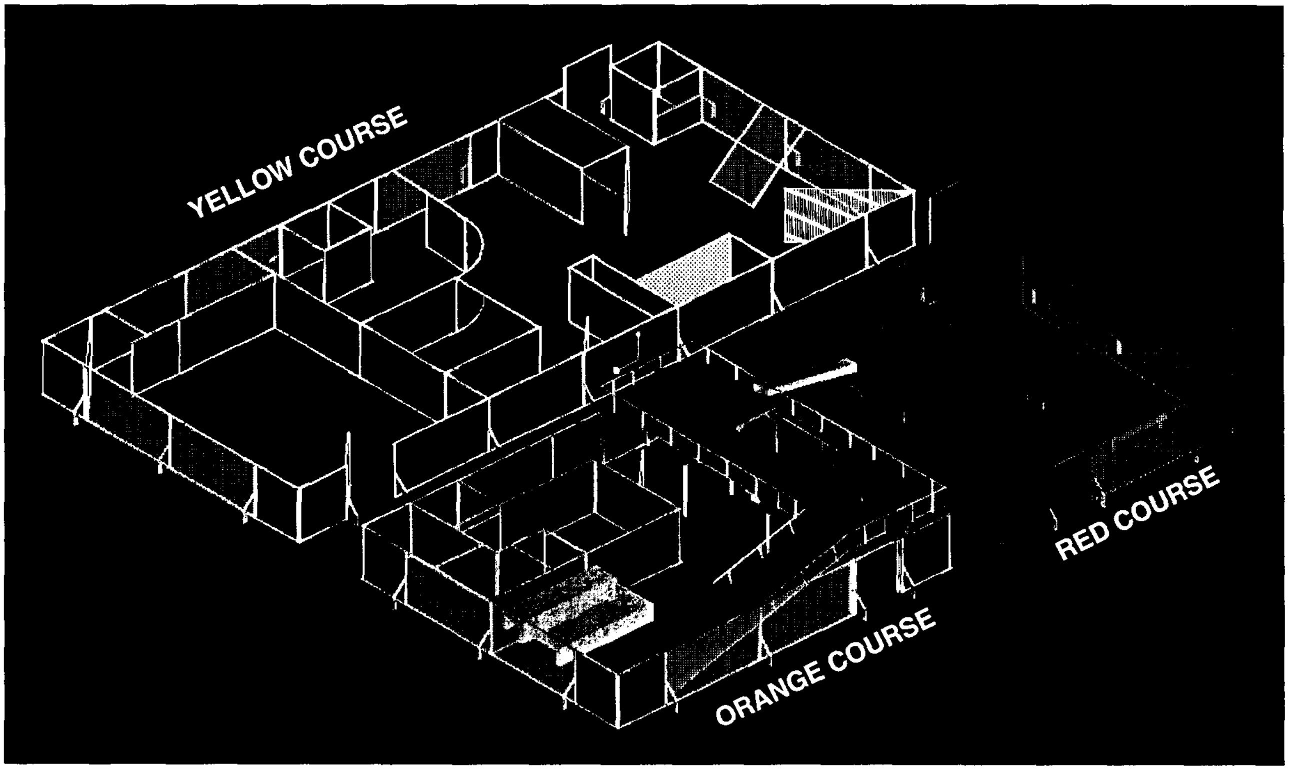

Figure 3: CAD View of the Test Course Layout

Yellow Course

Given the fact that participating teams, at least initially, are primarily from universities that may not have access to new agile robotic platforms, one design requirement was to have an area within the course where the mobility challenges are minimal. We call this area the "Yellow" course, and it is the easiest course to maneuver in. The floor of the Yellow course is flat and continuous. Passageways are wide enough to permit large robots, up to about 1 meter diameter, to pass easily.

Yet the Yellow course allows teams with sophisticated perception and planning to exercise their robots' capabilities. Some sensing challenges are as difficult in this section as in the others. There is highly reflective and highly absorbent material on the walls. Certain wall panels are clear Plexiglas, whereas others are covered in brightly patterned wallpaper. Some areas are dimly lit or accessible only from one direction. Victims are represented in all modalities (i.e., acoustically, visually, through motion, thermally, etc.) and are hidden from view under furnishings or in closed areas.

Orange Course

The Orange course is of intermediate difficulty. A second story is introduced extending the maze paradigm to three dimensions. And there are routes that only smaller robots may pass through. The robots have to climb stairs or a ramp in order to reach victims. Flooring materials of various kinds, such as carpeting, tile, and rubber, are introduced. Hazards, such as holes in the floor, exist. In order to be effective, the robot has to plan in a three-dimensional space. Larger robots are be able to navigate through some portions of this course, but not all.

Red Course

The Red course poses the most realistic representation of a collapsed structure. We do not anticipate that any of the contestants will be able to autonomously negotiate the Red course in the next couple of years. However, this section provides a performance goal for the teams to strive for. In the Red section, piles of rubble abound, lighting is minimal or non-existent in areas representing so-called "pancake collapses," and passageways are very narrow. The course is highly three-dimensional, from a mapping perspective. There are multiple floors which are surrounded by rubble piles that the robot has to traverse. All these features need to be mapped. Passageways under the rubble or through pipes have to be negotiated by the robots to reach certain areas or to get closer to victims. There are some portions of this course that can be traversed by the larger robots, but they are not be able to reach most of the victims. Larger robots would be best suited in marsupial configurations in this area.

Conclusion

An Urban Search and Rescue application for autonomous mobile robots poses several challenges that can be met only by highly intelligent, mobile systems. The variability, risk, and urgency inherent in USAR missions make this a good framework in which to begin measuring performance in controlled and reproducible situations. We believe that the test course we are developing can serve to elucidate performance measures for overall systems, as well as for components of intelligent systems.

Acknowledgments

The development of the USAR test course was carried out with the support from the Defense Advanced Research Agency (DARPA) Mobile Autonomous Robot Software Program and the Naval Research Laboratory.

References

Kitano, H. et al. 1999. RoboCup Rescue: Search and Rescue in Large-Scale Disasters as a Domain for Autonomous Agents Research. In Proceedings of the IEEE Conference on Man, Systems, and Cybernetics.

Blitch, J. 1996. Artificial Intelligence Technologies for Robot Assisted Urban Search and Rescue," Expert Systems with Applications, Vol. 11, No. 2, pp. 109-124, 1996.

Casper, J., and Murphy, R.R. 2000. Issues in Intelligent Robots for Search and Rescue. In proceedings of the SPIE Ground Vehicle Technology II Conference. Orlando, FL.