AI Magazine Volume 24 Number 1 (2003) (© AAAI)

The AAAI-2002 Robot Challenge

Benjamin Kuipers and Ashley Stroupe

■ The Eighteenth National Conference on Artificial Intelligence (AAAI-2002) Robot Challenge is part of an annual series of robot challenges and competitions. It is intended to promote the development of robot systems that interact intelligently with humans in natural environments. The Challenge task calls for a robot to attend the AAAI conference, which includes registering for the conference and giving a talk about itself. In this article, we review the task requirements, introduce the robots that participated at AAAI-2002 and describe the strengths and weaknesses of their performance.

The Eighteenth National Conference on Artificial Intelligence (AAAI-2002) Robot Challenge is part of an annual series of robot challenges and competitions. The purpose of the challenge is to promote the development of robot systems that interact intelligently with humans in natural environments. Additional goals of the challenge are to build multidisciplinary research groups to address these difficult problems and educate the public about the state of the art and challenges in current robotics research.

The Challenge task is to develop a robot that can attend the AAAI National Conference on Artificial Intelligence. This task involves many difficult subtasks, including navigating in a unknown crowded environment, locating a registration desk, interacting with other conference attendees, and giving a talk at a time and location provided at registration. The different subtasks require, or can benefit from, many different skills.

Task

Locating the registration area requires safe navigation within a crowded, dynamic environment. It can also require natural language processing and understanding for following directions, natural gesture understanding, sign reading, person following, elevator riding, and map building. Registration also requires natural language processing and understanding and can benefit from manipulation capabilities and the ability to stand in line. Interaction with attendees requires some natural language processing and understanding and can benefit from face recognition, badge reading, or some other form of identification and different types of interaction for different people and situations. Giving a talk requires navigation, interaction with a projector, and perhaps understanding content to give a better presentation and answer questions. Teams are ultimately expected to perform these functions completely autonomously without modification of the environment. At the present time, the level of autonomy is flexible, as is the ability to slightly modify the environment. This modification is permitted to allow research teams to focus on different aspects of the project at a time.

The particular course of this year's Challenge began in Edmonton, Canada, at the front door of the Shaw Conference Centre. The registration area is two floors down, which requires using an elevator. At the bottom of the elevator, the registration desk is not immediately visible (requiring directions, sign reading, or exploration). The registration desk is in a crowded environment (a condition greatly exacerbated during the Challenge because of the presence of onlookers). The talk was to be given directly following registration in the large Exhibition Hall. Robots were allowed to map the Exhibition Hall ahead of time, but maps were only made available for use once registration was completed. The talk took place in front of seating stands in the back of the Exhibition Hall, requiring navigation through the

Figure 1. IRobot: CoWorker in Edmonton Alberta's Shaw Conference Centre, AAAI–02.

crowded hall, the layout of which was changed frequently in small ways.

Three participants addressed different aspects of the whole challenge. The Massachusetts Institute of Technology (MIT) demonstrated autonomous mapping of the conference hall and rudimentary registration skills. iRobot Corporation demonstrated a mixed autonomous system, with safe-guarded teleoperation for navigation and televoice for interaction. The GRACE Team (Carnegie Mellon University [CMU], the Naval Research Laboratory [NRL], Metrica, Inc., Northwestern University, and Swarthmore College) attempted to perform all stages of the challenge autonomously, excluding interaction beyond that required for locating registration and registering.

iRobot: COWORKER

The COWORKER robot from iRobot is an internet controlled, wireless, mobile, remote telepresence platform (figure 1).[1] The user-friendly interface gives the human operator control over where the COWORKER goes, what it sees, and what it hears and provides an interface for speaking for interaction with others. The robot is equipped with sonars and a camera. There is even a laser pointer so that the user can highlight what he/she is referring to at the robot's location. In addition to joystick-type controls, the operator can specify a way point in the current field of view, and the robot moves autonomously to this way point, updating its location and avoiding obstacles.

During the COWORKER demonstration, the operator (Jim Allard of iRobot) remained out of the line of sight of the robot and relied only on the sensing of the robot. A human assistant pressed elevator buttons and provided feedback to the operator when asked. The COWORKER demonstration included starting at the upper level of the conference center, entering and exiting the elevator, and navigating to the registration desk. At the desk, the operator (through the interface) asked for registration materials to be placed on the robot. The robot was then navigated to the talk location. The talk was given by the operator.

The COWORKER system successfully and efficiently completed the navigation portions of the task. In comparison with GRACE (discussed later), which attempted the same task autonomously, COWORKER's success was owed to the presence of a human intelligence controlling the robot's behavior, which depends on a good interface that allows the human operator sufficient perceptual information to make appropriate decisions and high-level motor primitives (the specification of way points) to allow effective action.

Although iRobot's performance was impressive, it could have been improved with better capabilities for manipulation and interaction with bystanders. A manipulator capable of button pushing would certainly have been helpful for this particular task and seems more generally useful. Specifying a target for the button pusher could be done with the same type of user interface currently used to specify travel way points. Because COWORKER is a commercial product, and conference registration robotics is unlikely to be a major profit center, other considerations will probably determine whether our recommendations are followed.

Part of the benefit to the AI community of COWORKER's participation in the Robot Challenge is a clear comparison on the same task between an autonomous robot and one with human intelligence and knowledge in real-time control. Clearly, autonomous intelligent robotics still has quite a ways to go.

Massachusetts Institute of Technology: LEO

The MIT entry in the Challenge was from the lab of John Leonard in Ocean Engineering, who was assisted by Paul Newman and Michael Bosse. Their goal was to demonstrate autonomous, large-scale concurrent mapping and localization (CML). Autonomous navigation is a critical competence required for completion of the overall Challenge task. The ultimate goal is for a robot to arrive at the conference center with little or no a priori information and be able to find its way around and build a map that it can later use for efficient position estimation and path execution throughout the site.

This demonstration was based primarily on a set of new feature-based CML algorithms developed by the research group. The results from a real-time implementation of CML that was capable of autonomously returning to its start position, after a manually guided exploration stage, were presented at the 2002 International Conference on Robotics and Automation (Newman et al. 2002). For AAAI-2002, a new feature-based exploration strategy was developed to enable truly "hands-off" autonomous localization and mapping. The system was also integrated with a new technique known as ATLAS that enables CML to be performed efficiently in very large areas using a network of local maps (Leonard et al. 2003). Both methods work with either laser scanner or Polaroid sonar data.

The MIT team did not even bring its own robot to the conference. The complete system was integrated on an iRobot B21 (named LEO) that was made available to them for the demonstration by iRobot, Inc. LEO successfully operated at numerous locations in the Shaw Conference Centre for extended periods of time, including the street entrance to the Shaw Centre and the lobbies outside the conference presentation rooms. The system performed well, demonstrating the viability of a feature-based approach to CML using laser and/or sonar data in a large-scale, highly dynamic environment. Although LEO did not attempt to follow the script for entering the conference center, finding the registration booth, registering, and speaking, its mapping and localization abilities were impressive and are an essential part of the cognitive resources required for a robot to accomplish this task.

The MIT group would like to express its sincere appreciation to Doug Few, Laura Woodbury, and Jim Allard of iRobot, Inc., for their tremendous support during the conference.

GRACE

GRACE is the result of integrating separate research contributions from five different institutions: (1) CMU; (2) the NRL; (3) Metrica, Inc.; (4) Northwestern University; and (5) Swarthmore College. The project is under the direction of Reid Simmons (CMU) and Alan Schultz (NRL). The main goals of the GRACE team for this year were to (1) develop a framework in which to integrate software that had been developed by the various institutions onto a common hardware platform and (2) attempt to do the complete Challenge task autonomously, from beginning to end, to set a benchmark for future performance.

GRACE is an iRobot B21 with sonar sensors and a laser range finder (figures 2 and 3). It also has a Sony color camera with pan, tilt, and zoom; a separate stereo vision system mounted on a pan-tilt head (developed by Metrica); a wireless microphone; and a flat-panel display for showing GRACE's computer-animated face. NRL was responsible for multimodal interaction (speech and gesture). CMU was responsible for map-based navigation, registering (standing in line and interacting with the registrant), vision-based servoing to the registration desk, and elevator riding. Northwestern was responsible for having GRACE give a talk about itself, Metrica was responsible for stereo vision-based gesture recognition and people tracking, and Swarthmore was responsible for vision-based object tracking and reading of name tags. Overall integration was shared among the team members. A separate article in this issue describes the GRACE software system and architecture in more detail.

Anyone familiar with the difficulties of autonomous robotics in an unmodified environment, and with the difficulties of unifying the work of five different volunteer research groups, cannot be anything but vastly impressed with the achievement of the GRACE team. In spite of numerous difficulties along the way, GRACE completed all the subtasks, including delivering a presentation on its own technology. It attracted a great deal of attention, both from a large crowd of onlookers at the conference and from local and national media. Media coverage was generally effusive. For example, CNN said, "GRACE completed her task in just over an hour without any assistance from her human team." This was not entirely accurate, as we discuss in the detailed critique later.

The Robot Challenge is intended to be a task that will require several years to accomplish successfully. These critiques are intended as feedback to help guide future progress and to help educate the *AI Magazine* readership about the current state of autonomous robotics. They have already been communicated to the GRACE team, and some of these issues are discussed at greater length in the article about GRACE in this issue. With time to work before GRACE attempts the Robot Challenge at the Eighteenth International Joint Conference on Artificial Intelligence in Acapulco this year, it will certainly be possible to address some of these limitations, but there will surely be more to do after that.

Preparing the Environment

Although the intention was for the Challenge to take place in an unmodified environment, a few modifications were necessary. Walkways at the upper level of the conference center, and the elevator itself, were bounded by clear glass barriers, invisible to laser range finders and difficult for sonars to detect properly. Green tape was placed on the glass at the right level for GRACE to detect. We note that safety regulations often require that glass doors and walls be marked in ways that humans find easy to see to avoid human accidents. GRACE's sensors were at a different level than human sensors, so it needed special accommodation.

In addition, the registration desk was marked with a large pink poster to make it easy for GRACE to identify and move toward it from a distance. This, too, compensated (only partially) for GRACE's perceptual limitations and particularly for the difficulty of finding one of

Figure 2. Grace in the Shaw Conference Centre Hallway.

several registration booths at the edge of a large, complex open space.

In the future, the GRACE team might consider sensor fusion with both laser and sonar range finders to have a better chance of detecting even clear glass barriers. Improved visual and physical search methods might make finding the registration desk possible in a more realistic manner. Furthermore, the judges felt that this team might have missed an opportunity to educate the audience better about the significance of these features as illustrating the state of the art of robotics technology.

Communication to GRACE

The commercial speech-recognition interface (VIA VOICE) was clearly a weak link. GRACE's human assistant, Matt MacMahon, wore a microphone headset, but even so, communication was clearly a struggle, requiring statements to be repeated multiple times and corrected before being usable. In post mortem discussion,

Figure 3. GRACE *Approaching the AAAI–02 Registration Desk.*

the judges were told that the language understanding component worked correctly, once the speech input had been captured. Unfortunately, the failure of the front end made the language understanding difficult to see. Two gesture-recognition methods were planned, one using vision and one using Palm Pilot input, but neither was in working order during the Challenge itself, so they were not successfully demonstrated.

The speech input problems also made the registration desk interaction difficult. GRACE was following a script for the interaction but couldn't detect certain expected responses; so, it went into a strange, looping interaction pattern, repeating the same question at irregular intervals. Afterwards, the team was told of one of the inevitable integration screwups: The dictionary relevant to the registration desk interaction had somehow not been loaded at the correct time.

Communication from GRACE

The animated face on the LCD screen was a nice touch. Unfortunately, in the noisy environment of the convention center, the generated voice was often difficult to understand by many people. The words were displayed in "speech balloons," but these balloons were easy to miss. It is perhaps worth considering the kinds of very large-print displays used when Hollywood movies try to show computer interaction on screen.

Local Motion Control in Crowds

GRACE was almost always closely surrounded by a dense and eager crowd, pressing in on it from all sides. Even with assistance from the judges in clearing the way, it would often move a few feet, encounter an obstacle (often a person), stop, and ask whether it had reached the registration desk. Its human assistant would then struggle with the voice interface to explain that it had not reached the desk and orient it in the correct direction for further travel.

A more robust path-following control law would allow it to move more effectively through spaces in a crowd. A compass or ceiling camera would allow it to maintain its orientation while it avoided pedestrians. An interaction script designed for crowds would allow it to request encroaching people to move away. A robot registering for a conference is still unusual enough that a crowd can be reliably expected.

Error Recovery

There was an impressive bit of error recovery in the upstairs hallway that was observed and appreciated by the judges, although it might have been too subtle to be appreciated by the general audience. On its way to the elevator, GRACE oriented toward a patch of wall that it thought was the elevator door. Once there, it looked more carefully for the features expected of the elevator door, did not find them, and announced that it would look nearby for the elevator. It did so and was successful.

Standing in Line

Even including a dramatic misjudgment, GRACE's capability for standing in line was impressive. Once it had located the registration desk from its large pink sign, it approached the desk, identified people standing in line, and decided (incorrectly) where the end of the line was. Unfortunately, moving into place, it gave a solid body check to one of the judges (Leslie Kaelbling of MIT), which was not at all a wise move for a first-time registrant at a AAAI conference. It seems safe to assume that in the future, obstacle avoidance will be more effectively integrated into its routine for standing in line.

Localizing in the Map

In place of receiving a printed map of the convention center at the registration desk, Challenge participants were allowed to explore the environment in advance and build maps in their own internal representations. However, they were only allowed to access these maps after leaving the registration desk. (Visual comprehension of a graphic map is an interesting research problem, but it was not included as a requirement for the Robot Challenge.)

Unfortunately for GRACE, there were significant changes in the layout of the exhibition hall between the time it made its map and the time it had to find its way to the place for its lecture. It localized incorrectly, wandering into an exhibitor's booth when it believed it was following a path. Obstacle avoidance prevented further progress, and it was lost until it could be talked out and gotten back onto its route.

This problem is related to two serious research issues—(1) how to handle changing environments and (2) how to handle partially reliable map-based communication—so it certainly does not have a quick fix. In fact, the problem of incomplete or incorrect maps was included in a previous AAAI Robot Competition several years ago.

The Lecture

GRACE's lecture went quite well. It displayed a set of PowerPoint slides and gave a lecture on the methods it used in the Challenge. However, because the lecture was not interactive, it was likely difficult for the audience to tell that it was not simply following a script but was actually using phrases from the slides to retrieve fragments of explanation from a knowledge base and assembling them into its lecture on the spot.

Summary

Three judges (Leslie Kaelbling, Ben Kuipers, and Ben Wegbreit) evaluated the performance of the entries to the Robot Challenge. The awards were as follows:

Technical Award for Autonomous Localization and Mapping: MIT

Technical Award for Human-Computer Interaction: GRACE

Technical Award for Mixed Local Autonomy and Supervisory Control: iRobot

Technical Award for Robustness in Recovery from Action and Localization Errors: GRACE

One additional award was given as part of the overall AAAI Robot Competition event:

Ben Wegbreit Award for System Integration: GRACE

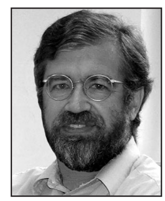
Note

1. www.irobot.com.

References

Leonard, J. J.; Rikoski, R. J.; Newman, P. M.; and Bosse, M. 2003. Mapping Partially Observable Features from Multiple Uncertain Vantage Points. *International Journal of Robotics Research.* Forthcoming.

Newman, P. M; Leonard, J. J.; Neira, J.; and Tardos, J. 2002. Explore and Return: Experimental Validation of Real-Time Concurrent Mapping and Localization. In Proceedings of the 2002 IEEE International Conference on Robotics and Automation, 1802–1809. Washington, D.C.: IEEE Computer Society.



Benjamin Kuipers holds an endowed professorship in computer sciences at the University of Texas at Austin. He investigates the representation of commonsense and expert knowledge, with particular emphasis on the effective use of incomplete knowledge. He received his B.A. from Swarthmore College and his Ph.D. from the Massachusetts Institute of Technology (MIT). He has held research or faculty appointments at MIT, Tufts University, and the University of Texas. His research accomplishments include developing the TOUR model of spatial knowledge in the cognitive map, the QSIM algorithm for qualitative simulation, access-limited logic for knowledge representation, and the spatial semantic hierarchy model of knowledge for robot exploration and mapping. He served as department chairperson from 1997 to 2001 and is a fellow of AAAI and IEEE.

Ashley W. Stroupe is a doctoral candidate at the Robotics Institute at Carnegie Mellon University. Her research focuses on the coordination of teams of behavior-based robots and multirobot localization, with applications including RoboCup. She received her B.S. from Harvey Mudd College, an M.S. from George Mason University, and an M.S. from Carnegie Mellon University.