AI Magazine Volume 23 Number 1 (2002) (© AAAI)

The Hors d'Oeuvres Event at the AAAI-2001 Mobile Robot Competition

Francois Michaud and David A. Gustafson

■ Serving hors d'oeuvres is not as easy as it might seem! You have to move carefully between people, gently and politely offer them hors d'oeuvres, make sure that you have not forgotten to serve someone in the room, and refill the serving tray when required. These are the challenges that robots have to face in the Hors d'Oeuvres, Anyone? event, part of the AAAI-2001 Mobile Robot Competition.

Hors d'Oeuvres, Anyone? For the fifth year that this event has now been held, five entries took on the challenge of creating service robots who can offer hors d'oeuvres to attendees of the robot exhibition. Such robots require the ability to move safely in a crowded environment, cover a serving area, find and stop at people to offer food and interact with them, detect when more food is needed, and take the actions necessary to refill the serving tray. Although the research topics addressed by this event are mostly targeted at human-robot interaction, one additional requirement this year was that the robots had to wear a server's badge to allow possible interactions with robots from other teams.

Interesting capabilities were demonstrated by all the entries. For example, Ron Nucci from the Seattle Robotics Club demonstrated a robot built completely with inexpensive off-the-shelf hardware (figure 1). The robot had voice-recognition capabilities using a demo program from Sensory Inc. called FLUENT SPEECH. A laser oriented toward the floor was able to detect people's feet, making the robot stop and start a dialogue with a potential customer.

The entry from Kansas State University consisted of two Pioneer 2-DX robots, a butler named CADBURY and a host named RICHIE, working as a team to serve the guests (figure 2). CADBURY seeks out guests and interacts with them to serve hors d'oeuvres, and RICHIE welcomes guests. CADBURY had a web camera and an infrared sensor mounted to provide perception and sense the presence of people by their temperature. RICHIE had only a web camera and used vision identify faces. Standard audio devices were connected to both robots. Mannequins were mounted on top of each robot to serve the twin purposes of giving the robots some color and increasing the range of their sensory devices. The robots communicated with each other through a local area network on wireless network cards on their laptop computers. Navigation was done using occupancy grids. The implemented strategy is as follows: CADBURY locates the closest obstacle in a room and investigates using the infrared capability to see if it is a human. If so, CADBURY asks for an order from a given set of hors d'oeuvres and records the response, which is analyzed by fast Fourier transforms to find a match with the expected responses. Once a match is found, CADBURY returns to the kitchen, fetches the ordered dish from the chef, returns to the guest, and serves him/her. CADBURY then moves on to find other guests. If a nonperson object is found, CADBURY tries to identify a robot badge on the object. On finding one, the other robot

Figure 1. Robot Entry from the Seattle Robotics Club.

is greeted, and a brief interaction follows. Otherwise, CADBURY surveys another quarter of the area. RICHIE was provided with a face-detection module, which filters non–skin-colored components from an image and tries to recognize known faces among the guests. On finding a known face, CADBURY is summoned, through a network message, to attend to the guest. If a new guest is found, the vision module on RICHIE tries to recognize the colored ribbon that hangs from the identification badge to identify specific people from the American Association for Artificial Intelligence. Because of technical difficulties, the robots did not perform well at the competition.

Swarthmore College, winner of the 2000 event (Maxwell et al. 2001), also participated in 2001. The primary innovation was reading the name tags of the conference attendees using a pan-tilt-zoom (PTZ) camera. This task involved three separate phases: (1) locating the name tags, (2) locating text on the name tags, and (3) identifying the individual letters (optical character recognition [OCR]). This problem turned out to be a challenge because (1) people move; (2) people don't always wear their name tags straight; and (3) even with a high-power zoom lens, the resolution captured by the video camera is, at best, half of what a typical scanner and OCR system uses. The robots use a simple state machine to identify name tags in a wide–field-of-view image using the color bars on a white field as a characteristic feature. The PTZ camera centers its image on a single name tag and then zooms in as quickly as possible, filling the image with the name tag (figure 3). At this point, the robot switches to a text identifier to put boxes around each line of text. The text identifier was based on the horizontal gradient-analysis technique used by Smith and Kanade (1998). While zoomed, the PTZ camera continues to center the text boxes using the text identifier. Once the system captures the text boxes, they are passed to a template-based optical character recognition system that integrates the matching with a name database and statistics on more than 2000 international names. When the same name is found in two of five recent detections, the robot speaks the person's name within a greeting. The badge reading system was integrated into the overall robot architecture, which is guided by a state machine. In addition to reading the name tags, these two robot entries could identify their refill station using red, white, and green flags; could identify skin color; could maneuver reasonably well in the crowd; and could identify when their tray was close to empty by tilting the PTZ camera down to look at it. Integration of these capabilities was provided by the REAPER architecture (Maxwell et al. 2001).

The Universidade de Aveiro in Portugal presented CARL, a fully autonomous robot developed to study the integration of communication, action, reasoning, and learning in robotics (CARL) (Seabra Lopes and Teixeira 2000). CARL (figure 4) is based on a Pioneer 2-DX mobile platform on which a fiberglass body, which includes a small tray, is mounted on top of the base, making CARL approximately 85 centimeters high. The fiberglass body carries a DA-400 v2 microphone array, a speaker, and a set of 10 infrared sensors for obstacle avoidance during navigation. Presence of food in the tray was also detected using infrared sensing. CARL communicates with human interactants by spoken language dialog. Speech recognition and synthesis is based on IBM VIAVOICE. Natural language parsing is done with the CPK natural language–processing suite. The grammar and the dialog manager support a representative set of speech acts (register, tell, ask, ask-if, achieve, and so on). The number of sentences recognized by the grammar was over 10,000. High-level reasoning, including inductive and deductive inference, was done in a PROLOG environment. Navigation is based on fusion of vision, sonar, and infrared sensing information. Based on a two-dimensional local map updated in real time, navigation is robust even in the presence of such "hard" obstacles as chairs and tables. Finally, CARL learns facts about the world from interaction with the human user. For example, it could learn that "Peter is in France." When it detects something on the scene that could be a person, it can also ask, "Is this a person?" The detected pattern, labeled with the obtained reply, is sent to CARL's "online learner" so that later, a person could be detected autonomously.

Last but not least, JOSÉ (figure 5) was the UBC Laboratory for Computational Intelligence's visually guided autonomous robot entry, a robotic waiter capable of navigating around a room populated by groups of people, politely serving appetizers. JOSÉ is an RWI B-14 robot with five video cameras as sensors: a three-camera stereo unit (Point Grey Research TRICLOPS), a Sony pan-tilt color camera, and a Logitech web cam. JOSÉ finds its way around by building a map (occupancy grid) of its environment using odometry and stereo data from the TRICLOPS camera system. Odometry errors can lead to significant discrepancies between actual and calculated locations. To deal with this problem, JOSÉ uses a localization technique based on the scale invariant feature transform (SIFT) (Lowe 1999). It builds a database of visual features in the environment, which it can then use to identify its position (Se, Lowe, and Little 2001). Simply put, JOSÉ can figure out where it is by recognizing the positions of things it has seen before. During the hors d'oeuvres competition, JOSÉ used its global localization abilities to find its base after running out of appetizers. JOSÉ also finds groups of people in a room by looking for clusters of skin-colored pixels through its color camera. The color images are registered with the stereo data, enabling location of the clusters in three dimensions. JOSÉ selects a cluster (group of people) to serve in an area that it has not previously visited. It navigates to the selected group, offers them appetizers, and then looks for another group. JOSÉ serves appetizers from a black tray monitored by the web cam. It knows that the tray is empty (and that it must return home) when the percentage of nonblack pixels in the web-cam images gets low. Finally, JOSÉ is dressed in a tuxedo complete with bow tie and white gloves. It is topped with an animated face on a screen that displays various facial expressions. JOSÉ also speaks in a pleasant tenor voice, politely offering food and making simple jokes.

As it turns out, JOSÉ won first prize in the event. It demonstrated its ability to find groups of people to serve, go back to a home base to pick up food, monitor the amount of food left on the tray, and present a polite and pleasant interface. JOSÉ's face, voice, and well-tailored

Figure 2. CADBURY and RICHIE, from Kansas State University.

Figure 3. Robot Entries from Swarthmore College.

Figure 4. CARL from the Universidade de Aveiro in Portugal.

dress were great crowd pleasers (it received the most votes from the public), evoking many smiles and laughter. People caught taking food from the tray without an offer were scolded by an angry looking JOSÉ. The guilty persons would surreptitiously glance around, then replace the stolen appetizer and apologize to JOSÉ. It is worth noting that JOSÉ can recognize other serving robots using SIFT features for object recognition. JOSÉ does not require any special markings to perform this task. However, the opportunity did not arise during the competition to fully test out this capability. Second place was a tie between Swarthmore and the Seattle Robotics Club. CARL received third place.

People really enjoyed interacting with the robots during the hors d'oeuvres event. Having food to serve helps people anticipate what kind of interactions they will experience with the robots. The robot servants rapidly become the center of attention unlike real people who serve food. The most difficult part of the serving task is the dynamic nature of groups of people, who are constantly moving and even interacting with the server robot in unanticipated ways (such as clustering around, asking questions, or purposefully foiling the robot's attempts to move), making navigation more difficult for the robots. Also, being in such noisy environmental conditions, it was hard for the teams to demonstrate vocal interactions with people. However, because the robots still have some limitations to overcome before becoming real servants, people focus on what the robots are trying to do as part of the interaction. Such problems have to be taken into consideration when we move robots out of the controlled conditions of research labs. The hors d'oeuvres event is a great opportunity to experiment in such context. The event also serves as a common evaluation test bed for the different approaches used to implement the intelligent decision-making processes required by the autonomous robot servants.

We are now thinking of extending the Hors d'Oeuvres, Anyone? event to have robots greet people entering the exhibition area, guide people to requested exhibition booths, and serve hors d'oeuvres in the exhibition area. Having such a variety of tasks will make it possible to emphasize human-robot interaction (in an entertaining fashion as well as possibly demonstrate vocal interaction in less noisy conditions) and localization in a crowded environment, with the abilities required to serve hors d'oeuvres. The name of the event might change, but the fundamental idea will be preserved. The contest will allow undergraduate students to get involved and will stimulate research activities of graduate students, make robotics visible to the AI community, and address the social role the robots should have and the expected impacts they have on the people.

Acknowledgments

The judges for the events were Harold Forbes, David Kortenkamp, and Francois Michaud. We would like to thank the teams that participated in this event as well as Bruce Maxwell of Swarthmore College, Luis Seabra Lopes of the Universidade de Aveiro in Portugal, Ron Nucci of the Seattle Robotics Club, Prabhu Raghunathan of Kansas State University, and Jesse Hoey of the University of British Columbia for providing the descriptions of their robot entries that were used to write this article.

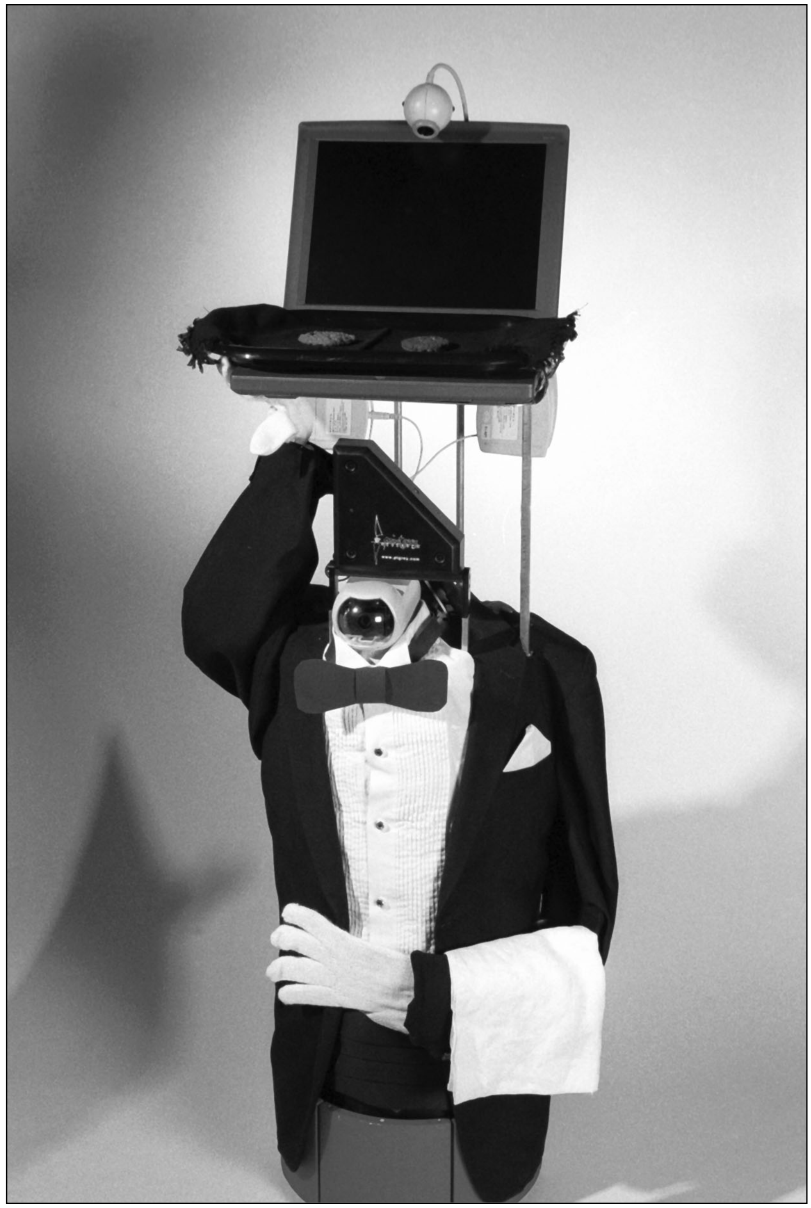


Figure 5. *JOSÉ from the University of British Columbia.*

References

Lowe, D. G. 1999. Object Recognition from Local Scale-Invariant Features. Paper presented at the 1999 International Conference on Computer Vision, 20–25 September, Corfu, Greece.

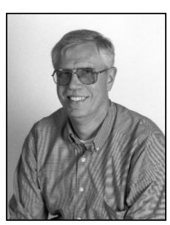
Maxwell, B. A.; Meeden, L. A.; Addo, N. S.; Dickson, P.; Fairfield, N.; Johnson, N.; Jones, E. G.; Kim, S.; Malla, P.; Murphy, M.; Rutter, R.; and Silk, E. 2001. REAPER: A Reflexive Architecture for Perceptive Agents. *AI Magazine* 22(1): 53–66.

Se, S.; Lowe, D. G.; and Little, J. J. 2001. Local and Global Localization for Mobile Robots Using Visual Landmarks. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, 414–420. Washington, D.C.: IEEE Computer Society.

Seabra Lopes, L., and Teixeira, A. J. S. 2000. Human-Robot Interaction through Spoken Language Dialogue. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Washington, D.C.: IEEE Computer Society.

Smith, M. A., and Kanade, T. 1998. Video Skimming and Characterization Using a Combination of Image and Language Understanding, In Proceedings of the IEEE International Workshop on Content-Based Access of Image and Video Databases, 61–70. Washington, D.C.: IEEE Computer Society.

Francois Michaud is the Canada research chairholder in autonomous mobile robots and intelligent systems and an associate professor in the Department of Electrical Engineering and Computer Engineering at the Université de Sherbrooke. He is the principal investigator for LABORIUS (www.gel.usherb.ca/laborius), a research laboratory on mobile robotics and intelligent systems working on applying AI methodologies in the design of intelligent autonomous systems that can assist humans in everyday lives. His research interests are architectural methodologies for intelligent decision making, autonomous mobile robotics, social robotics, robot learning, and intelligent systems. His e-mail address is francois.michaud@courrier.usherb.ca.



David A. Gustafson is a professor in the Computer and Information Sciences Department at Kansas State University. He received his Ph.D. from the University of Wisconsin at Madison. His research interests include software measurement, software engineering methodologies, validation techniques, AI techniques in software development, and software testing. His e-mail is address is dag@cis.ksu.edu.

AAAI-02
Edmonton/Alberta
IAAI-02

AAAI-2002 Robot Building Laboratory

Call for Participation

July 28–29, 2002

Collocated with AAAI-02 • Edmonton, Alberta, Canada

For years at AAAI conferences I've seen excited participants from the Robot Building Lab. Clearly they were having a terrific time. Finally at AAAI-98 I decided to treat myself and give it a try. It turned out to be the most fun I've had playing Legos since I was five! Now that I've actually helped build a robot, I no longer feel like so much of an imposter at Robotics Institute faculty meetings.

– Jack Mostow,
1998 AAI National Conference Cochair

The 2002 AAAI Robot Building Lab will be held July 28 - July 29, 2002, in Edmonton, Alberta, Canada. At this year's AAAI Robot Building Lab, participants will spend the day learning about how AI can (and can't) be integrated into the world of mobile robots. Most of the day will be hands on: building and programming small mobile robots to do a variety of tasks. Much of the current AI research deals with the actions of embedded agents. In this course it will become apparent that simulations of an agent's environment are often inadequate for effective evaluation of systems. The RBL will give the attendees the necessary information to start embedding their systems in physical agents – mobile robots that can interact with realistic environments.

Material Covered

- Realistic versus idealized robots
- Major components of robot systems
- Sensor and effector integration
- A crash course in behavior control programming
- Everything an AI researcher needs to know about PID control
- Vendors and suppliers for getting robots into your lab or home

This year's kit will include two processors, allow each group to create a duo of robots to work cooperatively. Functional mechanical modules will be available from the start. This year's kit will include 20 sensors (including sonar) and pneumatic actuators in addition to motors and servos. Participants will be able to spend their time designing and programming the robot, with only a bare minimum of LEGO hacking to get their robots to move reliably. (Plenty of LEGOs will be available for those who want to LEGO hack).

The RBL will be structured as follows. We'll begin with a brief tutorial on sensors, effectors, and robot capabilities to get everyone up to speed; then comes the actual robot building. Throughout the day there will be a series of short tutorials, both for individual teams and for the group as a whole, on particular aspects of robot building and programming.

The next day, all the robots will be completed and then displayed in the arena to show off their special capabilities and to compete head to head in a contest of speed and intelligence. This exhibition will be open to all conference attendees.

The goals of this lab event are as follows:

- Give all participants exposure to the intricacies of melding AI and robotics
- Show the value of performing AI experiments on physical devices
- Familiarize the participants with the current robotic experimental technology
- Give everyone a chance to play with AI that interacts with the physical world.

Target Audience

This lab is aimed at AI researchers, practitioners, and educators who want to move their systems out of simulations and into the physical world. A basic understanding of programming languages will be assumed.

Details

The RBL will take place all day Sunday and Monday July 28-29, 2002. This lab is being organized by AAAI and KISS Institute for Practical Robotics (www.kipr.org). For more information, send e-mail to rbl2002@kipr.org.

General Information

For RBL registration and for general information about AAAI-2002, please contact the American Association for Artificial Intelligence:

AAAI
445 Burgess Drive
Menlo Park, CA 94025
650-328-3123
650-321-4457 (fax)
ncai@aaai.org
www.aaai.org