The History of the Mobot Museum Robot Series: An Evolutionary Study

Thomas Willeke
Mobot, Inc.
Pittsburgh PA
twilleke@cs.stanford.edu

Clay Kunz
Mobot, Inc.
Pittsburgh PA
clay@cs.stanford.edu

Illah Nourbakhsh
Carnegie Mellon University
Pittsburgh PA
illah@cs.cmu.edu

Abstract:

This paper describes a long-term project to install socially interactive, autonomous mobile robots in public spaces. We have deployed four robots over the last three years, accumulating a total operational time of about six years. We introduce the robots, then focus on the lessons learned from one deployment to the next. The evolution of the robot-human interface is of particular interest, although other aspects of the robots' operations are briefly described.

Introduction

The history of autonomous mobile robotics research has largely been a story of closely supervised, isolated experiments on platforms which do not last long beyond the end of the experiment. In January 1998, we and others started work on Chips, an autonomous robot intended to be more than an experiment. Chips would become a permanent installation and member of the museum staff at the Carnegie Museum of Natural History in Pittsburgh, PA (Nourbakhsh et al. 1999).

Shortly thereafter, Mobot, Inc. was incorporated with a charter to improve and extend the Chips technology in a series of robot installations. Following Chips, three more robots have been developed in succession; three of the four still operate every day. Together, these robots have logged more than 2,000 total days of operation in their real-world public spaces.

In striving to deploy autonomous mobile robots in a social niche, we have two high-level goals. First, the robots must be autonomous to the largest extent possible. Human supervision of a full-time social robot is unacceptable. At most, the robots should only require occasional human help, and should request that help explicitly. Even the routine trip to the battery charger should be performed autonomously.

Second, since the robots would be deployed in public, they must have sufficiently rich personalities to achieve compelling and fruitful interaction with humans in their environments. A moving object without expressive interactivity would soon be moved into the closet.

We begin by presenting a brief overview of each of the four robots. Following this we discuss the evolution of our robot design in view of the goals of autonomy and personality.

Robot Overview

The four robots compared in this paper share the same operating system (RedHat Linux); the same robot platform (Nomadic Technologies XR4000); and the same programming environment (Gnu C++). The first robot, Chips, began work at the Carnegie Museum of Natural History on May 22, 1998. Chips operates exclusively in Dinosaur Hall, which contains large bone collections of *T. Rex* and other massive dinosaurs as well as ancillary exhibits focusing on topics such as paleogeology and ancient aquatic life. Chips's charter is to provide tours in Dinosaur Hall, presenting audiovisual information regarding both the large bone collections as well as the less frequented, smaller exhibits. Thus far Chips has been operating for almost 3 years, covering a total travel distance greater than 323 kilometers.

The second robot, Sweetlips, conducts tours in the Hall of North American Wildlife, also at the Carnegie Museum of Natural History (see Fig. 1). This space is composed of a number of dioramas in which preserved wildlife are shown in naturalistic settings. This portion of the museum has very low visitor traffic, so Sweetlips's charter is to both attract additional visitors to the Hall and to bring the static dioramas to life with high-quality footage of the same wildlife in their natural habitats. Sweetlips has been operating since April 1999, covering a total distance greater than 145 kilometers autonomously.

The third robot, Joe Historybot, operates in the atrium of the Heinz History Center. Its mission is to welcome visitors to this historical museum and provide both information and a tour of the atrium, which itself houses a number of significant exhibits. Joe provides historical information in an entertaining multimedia format. The robot also provides tutorials on speaking with a Pittsburgh accent and remotely triggers sound and light events associated with atrium exhibits. Joe has been operating since July 1999, covering a total distance greater than 130 kilometers autonomously.

Copyright © 2001, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

Figure 1: Sweetlips

The most recent Mobot robot, Adam 40-80, has operated in a variety of venues, including the Republican National Convention, the Democratic National Convention, a shopping mall, the National Aviary and most recently the Pittsburgh International Airport (see Fig. 2). Originally designed to promote Pittsburgh both statewide and nationally, Adam's charter is to engage passers-by with both information and challenges such as a trivia game. Instead of navigating a fixed tour route, Adam is also responsible for seeking out and approaching humans in order to engage them most efficiently. Adam has operated in a total of 6 venues for approximately 21 days.

Figure 2: Adam 40-80

An Evolutionary Study

The underlying goals of compelling interaction and maximal autonomy have remained constant throughout the creation of all four robots. However, each succeeding robot has been the product of a complete re-design based on lessons learned from the prior robots. Although some technical aspects have remained unchanged, such as the programming language and robot chassis, virtually all else has evolved in an effort to improve the autonomy and interactivity of the robots.

We are in the unique position of having an established trajectory of real-world interactive social robots. Studying the evolution of this robot series promises to uncover valuable information for the young science of social robotics. In the following two sections we discuss the evolution of the autonomy and interactivity of the Mobot robots.

On Robot Autonomy

The first requirement of a robot operating in a public space is safety, both for the general public and for the robot itself. At the heart of the matter is the robot's method for avoiding collisions, which must be especially robust, since the robots operate without supervision. It is notable that the collision avoidance code on these robots is by far the least changed over the course of their existence, confirming the functionality of the initial simple design. The robots use ultrasonic range-finding sensors (sonars) to detect obstacles, and move around them reactively, each cycle choosing the appropriate motion vector to take based strictly on the most recently available sensor data, along with restrictions on how far the robot is allowed to move out of its ideal (no-obstacle) trajectory. The code is extremely simple, with no explicit mapping or modeling of the world or of the sensors themselves. It is also easy to understand, and because of the lack of internal state, easy to debug (Nourbakhsh 2000). Because of the limited accuracy of sonar at close range, the robots will occasionally become stuck when they approach a wall too closely. Given the infrequency of this failure mode (less than twice per month), we feel the increased trust one can have in the robot's safety due to conservative motion to be worthwhile.

There is a great deal more to autonomy than safety. A robot must be able to interpret its own behavior, to determine whether or not it is functioning correctly. In order for humans to be confident in its ability to run without supervision, a robot must be able to determine on its own when a failure occurs. Early in the development of these platforms, we began using pagers, which the robot can signal via electronic mail. The ability to recognize failure and actively request help satisfies near-term requirements for autonomy. Of course the ultimate goal is that the robot never needs to send for help at all, so self-repair becomes a second step to self-diagnosis.

Initially, Chips sent for help quickly, giving up as soon as a failure was detected. Soon we began adding diagnostic methods to reset subsystems that weren't functioning correctly. This evolved into a general *method* for autonomy within our object-oriented architecture: every time a task is performed or an external piece of hardware is commanded, check the result for validity. If the result is invalid, reset the device or situation and try again.

When docking to recharge, for example, if the battery voltage fails to rise when the robot believes it is plugged in, the robot will reset the physical situation by backing out of the plug and into the hallway. Then, it will repeat the docking attempt. This "try again" policy is effective in robotics because, although the code is deterministic, there is sufficient nondeterminism in the environment that the same code may have different outcomes. We have further refined this policy with the caveat that the failure mode of an attempted task must be non-catastrophic for a retry to be possible.

The robots have evolved to make increasing use of this strategy, and now detect many abnormal situations, many of which are automatically corrected, including battery overcharging and undercharging, frame grabber anomalies, DVD player errors, bizarre encoder values (which would indicate that the robot had been pushed by an external force), emergency-stop activations, and the like.

Like other aspects of the robot, robot navigation and vision evolved over the four robots. Chips used a specific set of pink visual landmarks, one of which was three-dimensional, to provide corrections to simple encoder-only methods for determining location. As we installed robots in more locations, we added new kinds of visual landmarks, including sharp edges in intensity and rectangles of different color. We also added different methods for using them, allowing multiple landmarks to be tracked simultaneously (to deal with changing lighting conditions), and using landmarks to allow the robot to localize in more directions. We also used the same "try again" technique to make the landmark searching algorithm more robust. These changes are the subject of a companion paper being written concurrently.

On Human-Robot Interaction

Our second requirement was to deploy robots with compelling interactivity. As the science of Human-Robot Interaction (HRI) is in its infancy, it is not surprising that the robot interaction component was entirely redesigned in each subsequent deployment. Even so, we have reached several qualitative conclusions, which we will discuss here.

An interview with the exhibits maintenance staff of any large museum will drive home an important fact: people are basically destructive. Sometimes this is purposeful damage caused by malicious people. More frequently, curious individuals who are trying to better understand the robot cause damage. For example, some will attempt to push the robot off course to see if it will recover. Others will push any large red emergency stop button to see what happens.

Also, what attracts people varies greatly depending on the context of a particular public space. When in an "entertainment" space, such as a museum, people will be curious and attracted by new and unusual things. To that end the physical appearance of the robot is very important. But two other characteristics produce even better results: motion and awareness. When the robot is in motion, it draws the greatest attention from nearby people. To capitalize on this we made Adam twitch and move slightly while delivering longer presentations.

The most successful technique for attracting human attention is for a robot to demonstrate an awareness of human presence. Interactions between humans and complex machines are typically initiated by humans. When a robot deliberately faces a person and says "Hello," he or she is almost always both surprised and enthralled.

In contrast to entertainment venues, more utilitarian spaces such as shopping centers and office buildings elicit far less pronounced reactions. In these spaces people tend to have an agenda; they rush about and are less willing to be side-tracked by a new and entertaining creature. Early indications show that some success is possible using a much more socially aggressive robot which physically approaches individuals to initiate interaction.

In addition to attracting an audience, a robot must be able to retain one. Museum exhibit designers have tended to make their exhibits more interactive, often even taking on the characteristics of conversation. An exhibit may pose a question requiring the visitor to lift a panel or push a button to hear the correct answer. This is because attention tends to not stay focused through long presentations. By involving the visitors in the exhibit they stay more focused and curious about the information being conveyed.

We have found such techniques for retention to be equally valid for HRI (Nielsen 1993). Chips simply presents long (two minute) video clips at different locations throughout its tour path. As our robots evolved, so did their level of interactivity. Sweetlips includes the human observer in the process of choosing an appropriate tour theme. Joe goes further, answering many different classes of questions and even asking humans limited questions. Adam goes another step, playing trivia games with humans and taking polls. Such an exchange, where both the human and the robot can initiate the next part of the conversation, is essentially *dialogue*.

Because of a robot's particular sensory and effectory strengths, dialogue is multimodal and not necessarily verbal. Thus, while the human may be pushing buttons or using a touch screen, the robot may be responding with spoken words, music, graphics, video, text, physical gestures, and motion.

We learned several lessons from such robotic dialogue design. Firstly, there often will be a crowd of people around the robot, rather than a single person. Together with background noise from the environment, this makes it difficult for some people to hear the robot's responses if they are purely verbal. We therefore ensured that responses are always multimodal, including not only written screen text (e.g. captioning) but also graphics and video content.

Secondly, we found that long presentations, even movies, are guaranteed to drive the audience away. Instead, short responses combined with questions are most effective at extending the conversation. This parallels normal human interaction: the best conversations are dialogues between two people, not lectures. Finally, an aid to increasing the complexity of the dialogue is for the robot to have multiple ways of answering the same question so that it seems less scripted and more spontaneous, and therefore more interesting.

A final lesson learned with respect to HRI involves the psychological effect of creating an anthropomorphic robot. There are strong social rules governing appropriate behavior between humans (though these rules vary between cultures and segments of society), and there are other behavior patterns that people follow when interacting with machines and computers. A robot that looks somewhat human and has a rudimentary personality falls somewhere between these two modes.

The majority of people treat a robot as part human, part machine, clearly following some modified form of human interaction. Often they will treat the robot like a human by default, getting out of its way, and verbally responding to it. If they become interested in some feature of the robot, or want to investigate how it works, however, they will start treating it like a machine, ignoring its requests to move, and standing rudely in its way to see its reaction.

We believe humans use whichever social mode is most convenient for their short term goals. Fortunately, people will also often accommodate a robot that behaves in a fashion that would normally be unacceptable from another human. Since we were not actually in a position to do real social experiments (we had to keep our robot reasonably polite and could not experimentally find the boundary of unacceptable robot behavior) it is difficult to define the extent of this dynamic.

What we were able to experiment with is the robots' displays of emotional state. The main reason for a robot to display emotions is that humans expect and respond to them in somewhat predictable ways. People have a strong anthropomorphic urge and tend to attribute internal state to anything that behaves appropriately. People are also strongly conditioned to react to the emotions displayed by another person. These are powerful tendencies that robots should exploit.

These reactions are entirely behavioral. People cannot discern the true internal state of another human or robot. Their responses are thus entirely dependent upon perceived behavior. Chips and Sweetlips had sophisticated internal mood state machines that would change state over the course of the day, affecting the behavior of the robot. But since the visitors to the museum only interact with the robot for a short period of time, no one noticed these mood changes. Desigining Joe and Adam, we abandoned internal mood representation for a more transparent set of affective reactions to stimuli. On the other hand, if the robots were expected to interact with the same people on a daily basis, the internal moods would once again be useful.

As with the dialogue system, the richer the set of reactions the robot is capable of, the better. For instance, a good interaction model will greet humans in a variety of ways depending on context. If the robot is alone, it should be excited to see someone to interact with. Yet if the robot is busy giving a tour it should politely ask the person to join the tour or, failing that, to please get out of the way so that the tour group can move along.

Even more important than having reactions for all possible interaction contexts, it is critical that the robot's reactions are correct. If the robot begins talking to a wall or to thin air, it looks truly stupid. Just as moving safely through a crowd without hurting anyone is a basic required competence for a mobile robot, so total avoidance of stupid social interactions is a basic competence for a social robot. Generally, no one will notice if the robot fails to react to some indirect stimuli, but they will notice if the robot reacts inappropriately.

In summary, the interactivity of our robots has evolved along four axes: engagement, retention, dialogue, and anthropomorphic/affective qualities. Although this field of research is extremely young, it is already clear that there remains great pliability in the human-robot interaction model: human biases and bigotry regarding robots are not yet strong and fixed. We have an opportunity to design not just robot behavior, but the human behavior that will lead to the most fruitful possible human-robot interaction in the future.

Conclusion

Over the course of the last 2.5 years, we have built four robots, three of which operate on a daily basis with the public, autonomously and without human supervision. While this has been done before (Thrun et al. 2000, Sarcos 2001, Pyxis 2001), our robots are unique in their completely unsupervised free-roaming obstacle avoidance, and in their mission to entertain and inform the general public. We have learned many interesting lessons in attempting to meet the challenges described above; perhaps the most striking is that it actually is possible to deploy robots like these in the public over a long period of time. The robots described above are still running daily, and will hopefully continue to do so for an extended period of time.

In the course of watching the robots change, we have learned many lessons. First of all, it is important to make public robots resilient to physical abuse. People are not afraid to try to damage robots. In fact, they are eager to try to make them malfunction, and especially likely to press large red buttons to see what will happen. Children climb on, kick, and verbally abuse robots. Some fall in love with them. They must be able to handle all of these situations gracefully.

Secondly, when it comes to safety, simplicity in design and paranoia in implementation breeds confidence in deployment. Not surprisingly, once a good and easy to understand system is in place for collision avoidance, it tends not to change.

When robots are placed in public spaces, they must interact with people in such a way that will keep people's attention. The human robot interaction problem is in its infancy. While there have been many experiments in design, few of them have been deployed over the long term, to gauge general public acceptance. Our robots, even though they have been working for quite some time, only scratch the surface of experimentation in this domain. One initial conclusion is that a robot must have an adequate depth of dialog so that a human cannot immediately exhaust the robot's "conversation space," rendering the robot predictable, and therefore uninteresting. But in designing this personality, one must be as conservative as when designing obstacle avoidance code. Making obvious mistakes, such as talking to a potted plant, will cause the robot to be completely dismissed by the audience.

In the domain of autonomy, an approach to design and implementation that implicitly promotes fault-tolerance is important for the long-term survival of a robot. The basic "try again" approach works extremely well since the same code executed twice on a robotic platform will often yield different results. This approach, coupled with the ability of the robots to send pages when they need help, make human supervision refreshingly unnecessary. Even so, there are some types of failures that a robot cannot recover from completely, even if detection of the failure is possible. Drained batteries, burned-out fuses and lightbulbs, and cooked sonar transducers have brought each of the robots described down at various points in time, and the robots simply cannot fix themselves to that degree. Mobile robots still depend on humans for their continuing existence.

Acknowledgements

The original Natural History museum robot concept is due to Jay Apt and Red Whittaker. The following companies provided invaluable help throughout these projects: Nomadic Technologies, Magic Lantern, Maya Design, Redzone Robotics. Thanks also to co-creators of these robots: Judi Bobenage, Ben and Eva Carter, Vinton Coffman, Sebastien Grange, Iacovos Koumas, Ronald Lutz, Roland Meyer, Carolyn O'Brien, Todd Simonds, Alvaro Soto, David White.

References

Nielsen, J. *Usability Engineering*. Boston, MA: Academic Press. 1993.

Nourbakhsh, I., Bobenage, J., Grange, S., Lutz, R., Meyer, R. and Soto, A. An Affective Mobile Educator with a Full-time Job. *Artificial Intelligence*, 114 (1-2), pp. 95-124. October 1999.

Nourbakhsh, I. Property Mapping: A simple technique for mobile robot programming. In *Proceedings of AAAI 2000*. July 2000.

Pyxis Corporation. www.pyxis.com.

Sarcos Corporation. www.sarcos.com

Thrun, S., Beetz, M., Bennewitz, M., Burgard, W., Cremers, A.B., Dellaert, F., Fox, D., Haehnel, D., Rosenberg, C., Roy, N., Schulte, J., and Schulz, D. Probabilistic Algorithms and the Interactive Museum Tour-Guide Robot Minerva. *International Journal of Robotics Research,* 19(11):972-999, 2000.