Human-robot interaction: When robots want to play

Ashley Engelhardt

December 19, 2008

1 Introduction

Anyone who has ever owned a dog knows they can be relentless. When they want their human companion to play with them, dogs have the innate ability to solicit such behavior. What allows dogs to influence a person's behavior? In the context of robotics, this is an interesting issue. Understanding how dogs use non-verbal cues to get a desired reaction from a person could provide insight as to how emotions and emotional cues can be implemented on a robot.

Focusing specifically on play behavior, namely the non-human partner soliciting the human to play, I wanted to examine some of the factors that go into this type of interaction. There is little ethological research regarding the nature of human-dog play, despite the fact that play constitutes a large portion of the human-dog relationship [10]. Consequently, I chose to take a different approach and try to mimic dog-like behavior in a robot to examine how certain parameters influence the success of the human-robot interaction.

In this paper, I will first discuss the problem in a broad context and suggest possible approaches. Then I will describe my solution in detail, including an analysis of what worked and what did not work. Finally, I will discuss my unsuccessful attempt at setting up an experiment and relay insights that emerged from that failure.

2 Problem description

The human-dog relationship goes back more than 12,000 years [10]. The human-robotic dog relationship goes back a mere 20 years, but shows no sign of slowing down [2]. What is it about dogs that makes people drawn to them? One of the most endearing qualities of the domestic dog is their love of play. In fact, few species exhibit the type of interspecific play found between dog and owner and yet, very little research has been conducted to understand the nature of the play behavior [10].

A particularly interesting aspect of human-dog play behavior is that it can be initiated by either participant. The emergence of robotic pets such as Sony's AIBO and Ugobe's Pleo, which exhibit seemingly emotional qualities suggests that robotics may offer a unique perspective from which to examine play behavior. Namely, what factors influence the likelihood that a person will participate in a game initiated by his or her non-human companion?

A number of researchers have looked at the relationship between people and robotic dogs compared to the relationship between people and real dogs [6, 7, 3]. Kerepesi et al. [6] found that although there are differences in human-robot and human-dog interactions, there are similarities in the way in which people play with their companions. Likewise, Bartlett et al. [3] found that children overwhelmingly consider the Aibo a dog rather than a machine. These results indicate that there is an emotional component to robotic pets.

The idea of emotions in human-robot interaction has been explored in a variety of studies. Of particular interest for this project are those that deal with subtle affective behaviors, both intentional and unintentional [5, 8, 4]. Bruce et al. [5] examined the importance of having a robot face a person during conversation. Moshkina et al. [8] conducted a study the results of which suggest that people often associate emotion with visual cues given by a robot. Finally, Breazeal and Scassellati [4] discuss ways in which robots can indicate intentionality in the context of initiating an interaction with a human.

The previous studies served as inspiration for the present project, in which human-dog play behavior is used as a basis to study a subset of human-robot interaction. At a very basic level, the goal of this project is to examine the ability of a robot to solicit a particular behavior from human observers. A game of fetch was chosen as the intended behavior because it is familiar to most people and is often initiated by a non-human participant (i.e. a dog).

Before going into detail about my chosen approach, I will discuss some ways in which the project could have been completed. I make no claim that my approach is the best. Some decisions were based on available sensors, programming experience, and the simplicity allowed by a particular approach.

In terms of the overall architecture, a number of methods were considered. A subsumption architecture using obstacle detection at the lowest level with other, more specific layers on top is one option for a purely reactive architecture. Its main advantage is that the layered structure of the architecture makes it relatively easy to add layers later to expand the functionality of the robot. The drawback is that a purely reactive architecture does not allow for a hierarchical assignment of goals, which is necessary for cases in which one goal must be accomplished before another.

The study of behavior in an ethological context suggests the need for some form of planner [9]. Although animals are certainly reactive, there seems to be an underlying plan guiding their behavior. For that reason, a hybrid architecture consisting of a high-level planner with lower-level reactive layers can be used in an attempt to mimic animal behavior.

One key aspect of the interaction between robot and person is the robot's ability to correctly identify a person. Facial detection with a camera is one method for identifying a person, allowing the robot to distinguish between the front and back of a person. However, facial detection can have varying results based on lighting conditions and the angle of the face with respect to the camera. Likewise, using facial recognition discounts people who may be oriented facing the robot with their head turned in another direction. Another option is to use a laser to detect legs and assume that two legs found within a certain angle and distance represent a person. This method has the distinct disadvantage of having no way of distinguishing the front and back of a person and possibly mistaking other objects (poles, chairs, etc.) for people, but is advantageous in that it can provide a reliable location of the person with respect to the robot.

Another key aspect is how the robot recognizes the ball. As with person detection, the two obvious solutions involve either a camera or a laser. A camera can be used either for object recognition (i.e. recognize what a ball is) or color detection. Object recognition assumes the robot has some way of knowing what constitutes a ball. Given that a ball is a fairly distinguishable object, this may be an adequate method and does not restrict the ball to a certain color. Color detection does not rely on the robot's familiarity with the object of interest (the ball) but assumes the specified color is found only in the intended object. Alternatively, if mounted low enough, a laser can be used to detect a discontinuity and parameters can be established to determine if the discontinuity is (likely) a ball. Again, this is possible due to the consistent shape of the ball, although it would be difficult to distinguish between cylindrical and spherical objects.

In order to get the ball to the person, the robot needs to travel and have a way to determine how far it has moved. The simplest method is to use time and the programmed velocity to determine how far the robot has traveled or rotated. However, this method can be inaccurate if a reliable timer is not available. Using a technology such as GPS is another possibility, but requires the necessary equipment. Proprioception is also a valid approach, using internal information regarding tire size, gear ratios, and translational velocity to provide dead reckoning [1].

3 Solution

3.1 Description

For this project, the behaviors were tested first in a 2-dimensional simulated environment, then on the actual robot. It is important to note that both the real and simulated robots and environments had considerable limitations; as a result, the behaviors implemented differ in each. The robot used was the Pioneer 3-AT. The original four wheels were replaced with two smaller wheels on the front and a single castor on the back to allow for tighter turning. It was equipped with a 180° SICK laser for detecting obstacles and people (legs). A camera was mounted for color detection, specifically detection of the color red. A small red foam ball was used as the toy for the game of fetch. The ball was small enough that it did not interfere with the lasers. Figure 1 shows the physical configuration. In simulation, the robot had access to a vision server for camera access, a laser server, and a motion server for certain motor controls.

Figure 1: Pioneer used for demonstration

The robot was programmed to operate in an open room with few obstacles. In the physical environment, non-human obstacles were moved against the room walls so as to not interfere with the robot. The simulation did not include extraneous objects. The only item in the center of the room with the robot (real and simulated) was the red ball. For ease of detection, the ball was the only red object in the room.

I chose to implement a hybrid deliberative/reactive architecture. At a high level, there is a specific goal: solicit a person to play a game of fetch. This high-level goal is comprised of intermediate goals, such as detecting a ball, finding a person, etc. These intermediate goals and their relationships are represented in Figure 2. The reactive components of the architecture allow the

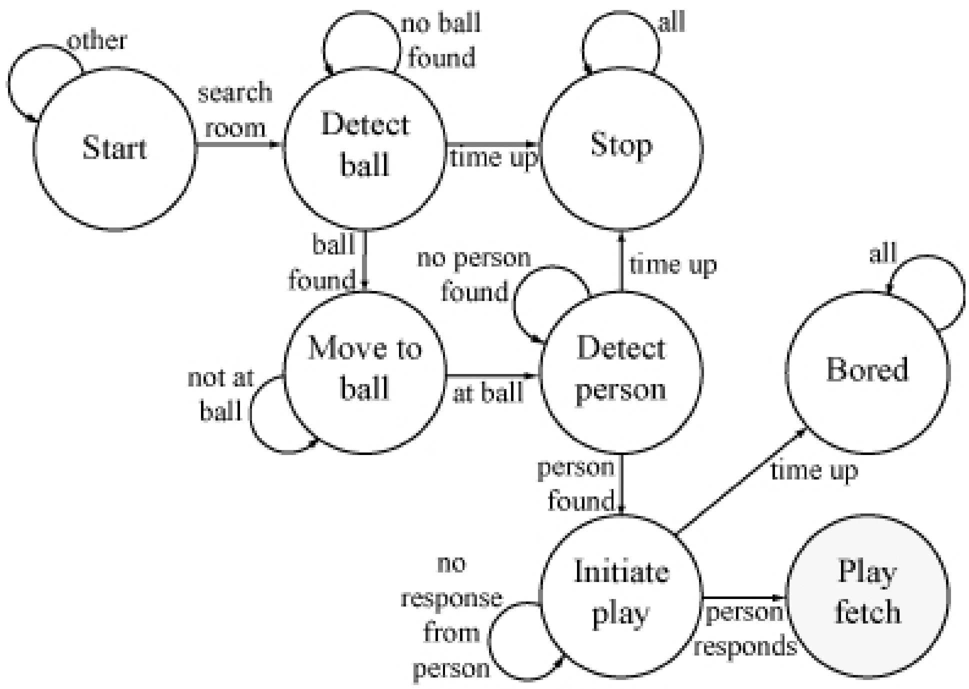

Figure 2: Finite state acceptor (FSA) diagram for play initiation - ”Fetch” behaviors are outside the scope of this project

robot to respond to unexpected events, such as the absence of a ball or person, the presence of an uncooperative person (i.e. not picking up the ball), or an unforeseen obstacle.

The first goal is to detect a red ball in the room. The robot is equipped with a camera used to detect the color red. During testing, the red ball is the only red object in the room; therefore the robot is programmed to assume the first red object detected is the ball. When a red object is detected, its angle with respect to the robot and area with respect to the entire camera frame are stored and used to navigate to the object. If no ball is detected, the robot sits for a while before searching again.

Once the ball is detected, the robot moves toward it. As the robot advances, the camera image is updated and the robot turns as needed to maintain a straight path to the ball. Since there is no straightforward method to extract the distance from the camera to the ball, the area is used instead. As the robot approaches the ball, it fills a larger portion of the camera frame. When the area reaches a certain threshold (determined through experimentation), the ball is directly in front of the robot.

The next objective is to find a person, which is accomplished using the laser to detect legs. Discontinuities in the laser readings are verified against a set of leg characteristics, and consecutive legs are then compared against a set of person characteristics (i.e. the legs must be within a certain distance and angle of one another). The robot follows a defined path, making a clockwise circle around the ball in search of a person. Although the laser is capable of detecting a person from 0 to 180°, the robot is programmed (in simulation) to only recognize a person between 80 and 90°. This limitation was imposed to make sure the person is roughly in front of the robot upon detection, thus making it much easier to align the person and the ball. In the physical environment,

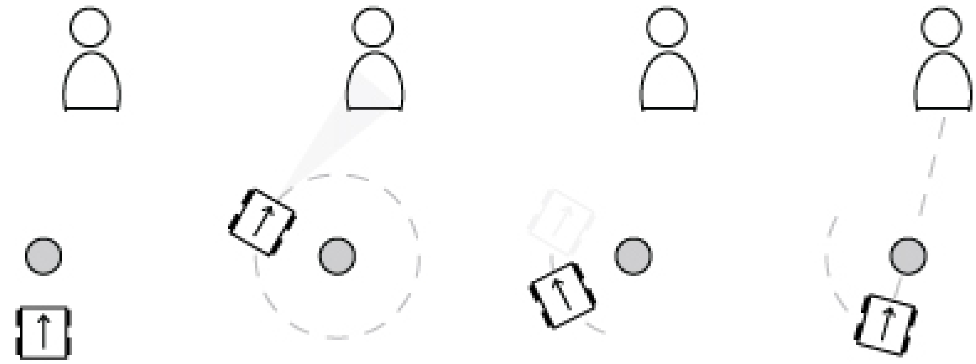

Figure 3: Process for lining up person, ball, and robot

the person detection range had to be increased do to delays in processing. Figure 3 depicts the alignment process. When a person is detected, the robot backtracks along a 90° arc and rotates 90° to the right. At that point, the robot, ball, and person are roughly aligned. The robot is then positioned to approach the person to attempt to engage in a game of fetch.

Enticing the detected person to play includes a number of steps in a prescribed order. Whether the robot goes through the entire process depends on the person's response. First, the robot pushes the ball toward the person and waits. If the person picks up the ball, the game has begun - the simulated robot drives off as if chasing the ball and the physical robot spins around. If the person does not respond, the robot moves to one side of the ball, turns slightly toward the person (i.e. looks at the person), turns back to the ball and backs up to the previous position. This sequence was intended to give the impression of "begging." The robot then pushes the ball closer, backs up, and waits again. As before, if the person picks up the ball, the robot has successfully started the game. However, if the person still does not respond after a predetermined amount of time, the robot gets bored. In its current implementation, "boredom" is indicated when the robot moves to a corner and sits.

An underlying obstacle detection layer allows the robot to avoid a collision as it navigates through the environment. Obstacle avoidance is achieved using discrete encodings, namely checking the right, left, and front for obstacles. This was particularly necessary to prevent the robot from running into a person standing near the ball.

3.2 Analysis

During the course of this project, some choices were made out of necessity (i.e. choice of robot) while others were intentional. The rationale behind the deliberative/reactive hybrid architecture is that neither approach fit on its own. The consistent, simplistic nature of the environment in which the robot is used suggested that a purely reactive architecture was not necessary. However, since the project focused on the interaction between the robot and a person, an inherently unpredictable relationship, a reactive, behavior-based component was needed for certain behaviors. Therefore, a hybrid approach made the most sense.

The main advantage of a hybrid approach is that it allowed me to constrain aspects unrelated to the human-robot interaction of interest. A significant constraint is the assumption that the ball is the only red object in the room. The color of the ball has no bearing on the interaction and can therefore be limited to a certain color without compromising the overall goal of the project. Likewise, choosing to use a red ball and eliminating other red objects from the environment has the practical advantage of eliminating the need for object recognition.

An important limitation of the hybrid architecture as used in this project is its lack of flexibility. The deliberative nature of the architecture mandates that the intermediate goals are completed in a specified order, at times creating unnecessary restrictions. For example, it is conceivable that the robot could initiate a game of fetch by first finding a person, then finding the ball. The rigid architecture specifies that the robot must find the ball before finding a person. Although it is admittedly limiting, the decision to find the ball before the person has a practical rationale. If the robot finds a person first, he or she could have moved by the time the robot reaches the ball.

Perhaps the most obviously planned aspect of the architecture involves the way in which the robot navigates around the ball as it searches for a person and aligns itself to push the ball. Circling the ball in a clockwise manner is not meant to mimic the actual behavior of a dog; rather, it is a consequence of the robot's inability to detect objects behind it. The circular path ensures that the ball remains next to, not behind, the robot. As the robot circles, the ball is out of camera range, but the knowledge that the ball is on the right allows the robot to safely back up along the path it traveled previously to align itself with the ball and the person. This type of expectation-based perception [1] may prove too restrictive to achieve a realistic ethological model, but was sufficient for the purpose of this project.

One aspect of the initial proposal that was not implemented is facial detection. The initial plan was that the robot would first look for a person facing in its direction then, if no one was found, it would search for a any person, regardless of where he or she was facing. The decision to drop facial detection was made due to time and logistical limitations. One camera was needed to detect the ball, so a second camera was required for facial detection. Mounting two cameras on the robot posed no problem, but it was unclear how difficult it would be to have two cameras accessing the same vision server simultaneously. Since facial detection was not critical to the overall project, it was eliminated.

Another piece of the original plan that was not accomplished is the set of experiments to determine how adjusting parameters such as approach speed, approach distance, etc. impacts the human in the human-robot interaction. Time was certainly a factor in forgoing the experiments, but a bigger issue was the difficulty translating the code from the simulated robot to the real robot. The simulated robot was able to use proprioception to determine how far it traveled and how far it turned based on gear ratios, velocity, and wheel size. Proprioception was not immediately available on the real robot due to the fact that it had different wheels than usual to allow for a tighter turning radius. As a result, all instances of turning or driving a specific distance had to be reprogrammed using for loops and counters to time the motion. Although this worked reasonably well for demonstration purposes, the rotations and movements were noticeably inaccurate and not suitable for testing.

Likewise, there were issues with the vision server on the real robot that made it impossible get the area of the ball from the camera image. The area was the only means for the robot to know when it was near the ball. As a temporary solution, the timing was reconfigured (again with for loops and counters) to give the appearance that the robot knew it was near the ball. In reality the robot could detect only the angle of the ball but had no way to discern the ball's distance.

This impacted the robot's ability to accurately move to the ball and line up with the person. Like proprioception, the inability to detect the area of the ball made the interaction between the robot and the ball significantly unstable.

3.3 Experiments and results

Unfortunately no experiments were conducted. The robot used for demonstration was notably inaccurate, a consequence of the use of counters to determine the timing. The absence of the ability to detect the area of the ball also hindered testing. Although no experiments were performed, the process of transferring the code to the physical robot revealed insights into areas that could be improved for better results.

First and foremost, the manner in which the robot circles the ball to find a person could be changed to a more natural motion. Using a different motion would enhance the impression that the robot is a spontaneous animal, not a programmed machine. Likewise, even though the robot did not resemble an animal, there was a feeling of being watched when the robot faced a person. That said, integrating longer periods for the robot to focus on (turn toward) a particular person before approaching him or her could again improve the robot's animal-like quality. Furthermore, increasing the number of times the robot turns to face the person throughout the play-initiation attempt could improve the interaction.

To improve the overall experiment, the robot could be programmed to track the ball. Currently the robot simply looks for the ball. If the ball is not where it was the last time the robot checked (assuming the robot has not moved), it is assumed the ball has been thrown and the game of fetch has begun. In reality, the ball could have rolled away and the person may in fact not be participating in a game of fetch. Adding the ability to track the ball could eliminate that limitation, providing a richer, more lifelike interaction.

The most notable improvement suggested while using the physical robot is the robot's ability to complete tasks in any order that makes sense. For single-session experiments, it may be sufficient to always have the robot find the ball before a person. However, for longer sessions, this could become monotonous. The robot should be programmed in a way that allows it to select the most appropriate sequence of tasks.

Regarding the types of experiments that could be performed, one possibility is to test the impact of changing the approach distance of the robot. As the robot approaches a person, there is a fine line between a friendly approach and a hostile approach. This is especially true if the person is not aware that the robot is trying to play. As I worked out the timing on the robot, there were occasions when the robot came right to my feet. It was apparent that such a close approach would likely be viewed as aggression and could deter future interaction with the robot. Determining the threshold for approach distance is one interesting aspect of human-robot interaction.

Other variables that could be tested are approach speed, persistence (i.e. how many times the robot tries to initiate play), and the force used to push the ball to the person. It would be particularly interesting to understand the types of emotions people associate with different robotic actions. For example, is a fast approach speed seen as excitement or anger?

4 Conclusion

With this project, my intention was to investigate how a robot can solicit a person to exhibit a particular behavior. A robot was designed in simulation to be able to initiate a game of fetch by first finding a red ball, navigating to the ball, finding a person, and bringing the ball to the person. The robot was programmed to try twice to get the person to pick up the ball, pushing the ball closer to the person with each attempt. If the person picks up the ball, the game begins. If not, the robot eventually gets bored and leaves.

Once the basic behaviors were implemented in simulation, the goal was to transfer the code to a physical robot and conduct experiments with participants. The experiments were meant to determine how adjusting various parameters (i.e. approach speed, approach distance, etc.) effects the robot's ability to entice the person to play fetch. However, a number of difficulties emerged when moving to the physical robot. The robot's inability to accurately travel a specified distance or angle and the inability to detect the area of the ball made the system's performance unpredictable. As a result, experiments were not performed.

In the future, before experiments are performed, a number of changes are suggested. First and foremost the robot needs a way to travel a certain distance consistently and detect the distance to the ball. Also, the method the robot uses to approach the ball should be modified to resemble a more natural behavior (i.e. not circling the ball). Once these changes are made, the resulting interaction between robot and human will be more indicative of natural human-dog interaction.

5 References

[1] ARKIN, R. *Behavior-Based Robotics*. MIT Press, Cambridge, MA, 1998.

[2] ARKIN, R., FUJITA, M., TAKAGI, T., AND HASEGAWA, R. Ethological modeling and architecture for an entertainment robot. In *Mobile Robot Laboratory Publications*. Georgia Institute of Technology, 2001.

[3] BARTLETT, B., ESTIVILL-CASTRO, V., AND SEYNOM, S. Dogs or robots: Why do children see them as robotic pets rather than canine machines? In *ACM International Conference Proceedings Series* (2004), vol. 53, pp. 7–14.

[4] BREAZEAL, C., AND SCASSELLATI, B. How to build robots that make friends and influence people. *Intelligent robots and systems 2* (1999), 858–863.

[5] BRUCE, A., NOURBAKHSH, I., AND SIMMONS, R. The role of expressiveness and attention in human-robot interaction. *Robotics and Automation 4* (2002), 4138–4142.

[6] KEREPESI, A., JONSSON, G., KUBINYI, E., AND MIKLÒSI, A. Can robots replace dogs? comparison of temporal patterns in dog-human and robot-human interactions. In *Human-Robot Interaction*, N. Sarkar, Ed. Itech Education and Publishing, Vienna, Austria, 2007, pp. 201–212.

[7] MELSON, G., KAHN, P., BECK, A., FRIEDMAN, B., ROBERTS, T., AND GARRETT, E. Robots as dogs? - children's interactions with the robotic dog aibo and a live australian shepherd. In *CHI 2005* (2005), pp. 1649–1652.

[8] Moshkina, L., and Arkin, R. Human perspective on affective robotic behavior: A longitudinal study. *Intelligent robots and systems* (2005), 1444–1451.

[9] Pfeifer, R., and Scheier, C. *Understanding Intelligence*. MIT Press, Cambridge, MA, 2001.

[10] Rooney, N., Bradshaw, J., and Robinson, I. A comparison of dog-dog and dog-human play behaviour. *Applied Animal Behavior Sciences 66*, 3 (1999), 235–248.