MAS 965 Relational Machines: Final Project Report

Andrew Brooks

May 16, 2005

1 Overview

The purpose of this project was to design a demonstration system that used an expressive robot assistant to encourage human patients to perform physical therapy exercises. The style of the relationship was one of a collaborative partnership. Through teamwork, the human and the robot labor together to achieve better physical performance for both of them. The goals of the system were thus to motivate the exercise itself, to provide natural and compelling performance feedback, and to be generally entertaining and empowering for the human.

2 Early User Feedback

I interviewed three individuals concerning this project. One was a researcher in the field of technology-assisted physical therapy, one was a product designer of physical therapy assistive devices for children, and one was a young sufferer of a recent stroke and associated physical therapy.

The researcher provided high-level guidance for the design of this type of system. He pointed out that there is a need for simultaneous feedback of the human's performance. This inspired the use of the robot's limbs to redisplay the human's performance while the face displays an affective rendition of the performance. The researcher agreed that displaying the human's motion on the robot can be more direct and easy to follow than a graphical representation. He also mentioned that the interaction itself must be kept simple so that patients can understand it even in the presence of complicating conditions such as visual neglect. Hence the movements and goals have been limited to iconic, easy to understand states. Finally he noted that tasks should have a scale of challenge levels, to provide more granularity of visible improvement.

The product designer provided examples of her company's current products which perform this sort of function. The interaction that they have designed is quite simple. The user selects from a number of activities that they can assist a graphical avatar in performing, such as blowing bubbles. Then the performance of the avatar matches very basic action scales of the patients, such as how hard they press a button. She appreciated the concept of this project with the physical robot. She pointed out that motions should be kept simple, such as how far a joint could be extended, or how high an arm could be lifted - in this fashion mirroring the button scenario. Her product did not have a relational, memory aspect to it, and she was in favour of this concept. She emphasized that it was important to have an interaction that was empowering for the patient and that the patient would feel proud of using.

The stroke patient made a number of useful observations from the point of view of the end user. Although she had suffered partial facial paralysis, she was not given any face exercises. She found that she did not do most of the exercises she had been given - for example, she did not palpate a squishy ball she had been given to exercise her hand with, because she did not feel it accomplished anything in addition to the exercise. Instead, she would perform minimal tasks while doing other things, for example squeezing the edge of the table while eating or talking, or manipulating the cards while playing poker. She liked the idea of a game interaction that kept track of performance, although she said she had not played video games since her stroke since she found her hand's reaction time could not keep up with what was necessary. She also pointed out that her improvement in motion quality and range of movement was not a linear progression, but came and went to some degree. She said that her touch sensitivity was also degraded, and that the therapists had not thought of any ideas for exercises to improve this, and that in general the state of her performance was hard for her to judge, even with gross motion capabilities such as shoulder movement. She said that trembling and fluidity of motion would be useful things to get better feedback of.

3 System Description

The implemented system consisted of the following components:

- A visual sensor of human arm pose. The system used was VTracker, developed by Dr David Demirdjian of MIT CSAIL. The sensor maps a prototypical articulated humanoid to stereo range data, and returns the 3-D coordinate positions of the wrist, elbow and shoulder for each arm. I implemented a remote receiver for this data that converted these positions into joint-angle data for a prototypical humanoid having a similar skeleton to Leonardo, our humanoid robot.

- An imitative body mapping system that allows the robot to re-express the human's movements. This is an imitation in which the robot babbles through its motor repertoire while a human (potentially the therapist) imitates the robot's poses. The robot assumes compliance on the part of the human trainer, and that the correspondence between the pairs of poses will be essentially good. The 3-D coordinate position of the robot's right shoulder, elbow and wrist and the information from the body tracker are used to train a linear model between these coordinate spaces. During the post-training phase, when the robot imitates the human, the position of the human's joints can be mapped to the desired position of the robot's joints. These positions are then fed into the kinematic solver, which produces joint-angle data that the robot can use to match the pose.

- An exercise representation. The desired exercises are stored as a sequence of arm poses set by the therapost. Although the exercise poses are set as complete arm poses, only the most distal position (the wrist) is used for comparison. The exercise is then set up as a finite state machine (FSM), where the occupancy of a given state is defined by having the hand within a thresholded linear distance of the desired hand position. The FSM is monotonic, in that from each state there can only be one next state.

- A representation of human performance. Human performance is stored as the times between reaching adjacent states, and the affective state of the robot at the end of each full state cycle, represented as a floating point value between zero and one. From this data can be extracted

Figure 1: A human subject is visually tracked as she performs the exercise with the simulated robot companion.

the number of complete exercise cycles performed and the rapidity of individual state cycles, and any potential correlation between this measure and the affective state of the robot. Each time an exercise cycle is completed, the robot's affective state value increases by a value δ_h. At any timestep at which a cycle is not completed, the robot's affective state value is reduced by a value δ_{s1}, unless the human's arm is in a position indicating relaxation. If this is the case, and a time interval δ_t since completion of the most recent cycle is exceeded, the robot's affective state decreases instead by δ_{s2}, where $\delta_{s2} >> \delta_{s1}$.

- An affect-based feedback channel from the robot to the human. The robot is outfitted with two custom poses representing extremes of affect: very sad and very happy. The robot's affective state at a given time is then used to produce a linear blend between these two poses. This allows the human to receive instantaneous feedback as to the satisfactory or unsatisfactory progress of the team's exercise task in a natural and intuitive way. Furthermore, once the robot reaches within δ_h of its maximum affective state, the robot indicates to the human with a satisfied head-nodding animation that the current exercise task is complete. This gives the human a fixed goal for which to strive.

4 Interaction

The demonstration system proceeded with the following interaction:

The patient approaches the robot, and the vision system begins tracking him or her. The robot demonstrates the exercise task by running through the poses set by the therapist. The robot then begins mimicking the human's arm pose and adopts a neutral affective state. The human's action (or inaction) then begins to modify the robot's affective state as described above. For the purposes of this demonstration, the exercise was a two-state task consisting of raising the arm above the

Figure 2: The simulated robot companion uses its affective display capabilities to give feedback on the state of the shared exercise task as it imitates the patient's arm movement.

shoulder and then lowering it below the waist.

5 Results

Due to concerns about the enforcement of limits on the robot's movement, for this project the "robot" used was the kinematically identical graphical simulation of the real robot, "Virtual Leo", rather than the robot itself. The simulation was projected at a large size on an LCD computer monitor in front of the patient.

One naive demonstration subject was allowed to participate in the interaction, a healthy female MIT student aged 21 years. She had no prior experience interacting with a real or virtual humanoid robot nor a computer vision system. It was explained to her verbally that she would be perfoming an exercise task in which she should copy the action that the robot demonstrated to her, and that she and the robot would be working together and her goal should be to make the robot happy. The vision system and the robot's body mapping had previously been calibrated by the experimenter, who is about 12 inches taller than the subject. Nevertheless, the system exhibited more than adequate mapping of the subject.

The subject appeared to be enthused with the prospect of making the robot happy, and performed the task with vigor. Compliance of the subject was therefore not an issue. The subject successfully completed the exercise up to the robot's exercise termination point. Following this, the subject stated that she had received a "good workout" and that the robot's happiness had motivated her to continue the exercises. The subject and the image of the robot on the screen were alternately videotaped during the interaction.

6 Discussion and Conclusions

In general the demonstration appeared to work effectively and was theoretically easy for a non-roboticist to set up and calibrate, and to explain to a patient. The issue of patient compliance was only weakly dealt with, however. The task is simple enough that there is really no way for the patient to "trick" the system while not performing the exercise, and in the case of non-compliance the robot becomes disaffected much more quickly, but there is no reason to expect that a truly non-compliant patient will care about the robot's happiness. The system would be equipped to report this non-compliance back to the therapist. However it may be wise to approach such a scenario with care. If patients believe that the robot is "spying" on them, they may be even more disinclined to cooperate with it. It would be better to make the task more engaging (for example, having multiple ancillary tasks that the exercise is designed to achieve, such as popping virtual bubbles or picking virtual cherries from a tree) with rewards of finding out what the next task is, rather than seeking to punish non-compliance — a carrot rather than stick approach.

The use of the maximum emotional state to terminate the exercise also gave rise to concerns. For example, if the patient proved unable to reach this state, would it be a source of frustration? On the other hand, if the patient is always able to reach this state, then the robot in a sense becomes much less "relational" — if each session ends with the robot extremely happy, then there is no sense of progression over time, or feeling of picking up the next session where you left off. It seems that having a more complex emotional model for the robot, and a mixture of reward states (e.g. rewarding a successful session with a story, conversation or skit routine performed by the robot) could be a much better long-term solution.

Since the work was performed on the simulated robot, it remains an open question of whether the real robot would have made the experience significantly different. On one hand, the robot makes it easier to visualize the motion, but this may be negated by the fact that noise in the vision tracking makes it difficult to precisely replicate the quality of the human's motion, and due to the idiosyncratic design of Leonardo's shoulder (one of the shoulder joints must undergo a complete reversal in order to transition the hand from the lower hemisphere of motion to the upper hemisphere) it is not possible for the robot to exactly replicate the human's movement for the arm lifting task. On the other hand, the fact that Leonardo's face is not yet attached means that the real robot has reduced affective display capabilities than the simulated robot. In conclusion, I believe that the simulated robot was the correct choice for the circumstances of this demonstration.

To summarize, this appeared to be an adequate proof of the concept of using a relational machine to mediate a physical therapy task. Future work must concentrate on giving the robot more facilities for relating to the human, in particular more engaging task modes and more mechanisms for rewarding the patient for compliance and persistence.