Figure 1: RoboPal, a local and video conferencing robot

RoboPal: A Video and Local Conferencing Interaction Robot

Eric Berg
eb645@cornell.edu
Cornell University
Ithaca, NY

Nicole Lin
nl392@cornell.edu
Cornell University
Ithaca, NY

Michael Xiao
mfx2@cornell.edu
Cornell University
Ithaca, NY

Adam Weld
aw698@cornell.edu
Cornell University
Ithaca, NY

Ho-Jung Yang
hy379@cornell.edu
Cornell University
Ithaca, NY

ABSTRACT

We present the ideation process, design, and results behind RoboPal, a robot aimed to make local and video conference meetings more productive, personal, and inviting. The robot does this by moving the camera, a clipped smartphone, with its five degrees of freedom, shifting its center of attention as the conversation commences. When single users are speaking, the robot is able to spotlight them such that others will listen and provide backchanneling throughout the conversation. In situations in which the conversation becomes unproductive (ie. off topic or argument), RoboPal is able to draw the group's attention and steer the meeting back on track. To gather input and reactions on the use of the robot, we conducted several sessions with the island survival challenge and found that though some users felt that RoboPal could be distracting or even intimidating, there is great potential for the robot to make users feel heard and improve teamwork among small groups.

KEYWORDS

HRI, human-robot interaction, social robotics

Figure 2: Example of a crowded meeting frame

INTRODUCTION

The scope of this project addresses a robotic based approach to facilitate video and local conferences. Often times in large video conferences, cameras are wide angled and unfocused, with many participants attempting to squeeze themselves into the frame. When this happens, the participants are more likely to feel disconnected to the meeting, as it gets less personal and engaging, especially in longer meetings. This results in unproductive meetings which can be an inefficient use of many people's time.

We aimed to address this problem by designing a robot that would make these video conferencing situations more personal. RoboPal is designed after a desktop lamp, a common appliance often interacted with. The head of the lamp can clip in a phone while the conference call is happening such that the robot is free to steer the presenter on camera to the speaker or center of attention in the room. To do this, RoboPal is equipped with a stepper motor to move the base, two large servo motors to control the main arm joints, and two micro servos to control the head movements. With these five degrees of freedom, RoboPal is able to not only move the camera around but also perform actions such as nodding, shaking its head, and acting interested or afraid. We took advantage of these peripheral features to also facilitate behavior of the participants of the meeting. For example, RoboPal is designed to be able to encourage a quiet speaker to invite their input or calm down an argument between two members. Unlike most other work in the HRI domain, RoboPal is built around supporting human to human interaction rather than direct human to robot interaction. In this paper, we highlight our ideation process for designing both the behavioral and mechanical portions of RoboPal, including the walkthrough of how we made our design decisions. We then summarize our results and findings after conducting the Island Survival Testing and draw some insightful conclusions about the effectiveness and potential of RoboPal.

RELATED WORK

A. Engagement

Research has shown that group's "collective intelligence" is closely tied to group dynamics rather than individual abilities of members [8]. Rather than assembling "dream teams," the psychological safety and structure of a group is a better indicator of a successful, productive team [6]. Though these studies are mostly of human groups, the results have inspired research examining if social robots creating a positive environment and facilitating group work could produce similar results. The work by Tennent demonstrates that even non-humanoid robots encouraging participant engagement through its movement and implicit interactions increased group problem solving performance and back-channeling [7].

Figure 3: Early storyboards for RoboPal demonstrating its behaviour in a group meeting

B. Conversational Gaze

Gaze cues have been shown to communicate intent even in human-robot interactions. In Moon's experiments looking into the reaction time of handing over a water bottle from a robot to human participants, shared attention gazes increased reaction times and yielded positive perceptions of the robot [4]. The study of gaze cues from robots extends to conversations also. While some settings might have fixed roles in conversation (e.g lectures), most conversations have different speakers at different times. Research has shown that role signaling cues are generally complied with in these settings and feelings of group cohesion (including the robot) are stronger when robots help facilitate conversations [5].

C. Anthropomorphic Design of Social Robots

The appearance and function of a robot affects the way people interact with it. Creating extremely humanoid robots may cause users to create a mental model of the robot too high to reach. However, a human-like appearance on a robot is shown to increase engagement and perception of the robotâĂŹs dominance, trustworthiness, responsiveness, and respectability [1]. Social robots with human-like characteristics and movement like Hoffman's Kip1 elicit empathy from participants in the field [2]. Instead of relying on robots explicitly spelling out how to interact with it, using nonverbal human-like movements can express intent through implicit interactions [3].

MOTIVATION

In beginning to generate a concept for a robot that supports work in groups, we started to brainstorm both objects that could be roboticized and human-robot interactions that could be streamlined or improved. In this initial brainstorming session, many of our ideas revolved around everyday objects (tables, chairs, clothes, and appliances) and more familiar problem spaces (office, school, art, travelling, kids, retail). We came up with several thoroughly developed ideas for robots before discussing which one to focus on. These included a set of self arranging furniture, watches to sync meeting times, a tie that could help a user remember new names in the office, and an auto-sorting toolbox. The robot lamp idea stood out to us among these other good ideas due to its feasibility to build within the given eight-week time frame, its test-ability among our peers (i.e. we would not have to find external groups to conduct testing), and is mechanical complexity. Furthermore, each group member had a plethora of visions and features to add, highlighting that we were passionate about delivering a product like RoboPal. The lamp itself was also an enticing object to roboticize as it is a common inanimate object in group meetings and would not evoke human-level expectations of its performance. They are already used on bigger stages like theatres or lecture halls to spotlight speakers and this movement would pair nicely as a video conferencing feature. This ability to spotlight group members stemmed from how easily a lamp's body could mimic human actions to perform conversational actions.

After committing to RoboPal, we wanted more clearly define its goals before moving directly into the design phase. We did this by creating storyboards to highlight what its use would look like and how it would be used. A couple storyboards shown in Figure 3 were constructed to envision RoboPalâĂŹs use in multi-group conferences talking with one remote member.

These early storyboards explore RoboPal's methods of facilitating group work. On the left (Figure 3.a.1), a meeting starts with RoboPal waving its head up and down to signal the start of a meeting. When a person starts talking (Figure 3.a.2), RoboPal spotlights them and then pans around the room when no one has anything else to say (Figure 3.a.3). When an argument starts, RoboPal's head droops down and its lights fade in and out (Figure 3.a.4). The storyboard on the right begins with the same premise where group work is given and then RoboPal extends its body to focus on a member who begins talking (Figure 3.b.2). When two people have a discussion, RoboPal backs up to include both members in the frame of the video (Figure 3.b.3). Once the discussion explodes into an argument, RoboPal shakes its head to catch the attention of the arguers and encourages the video conferencing member to call attention to the chaos (Figure 3.b.4).

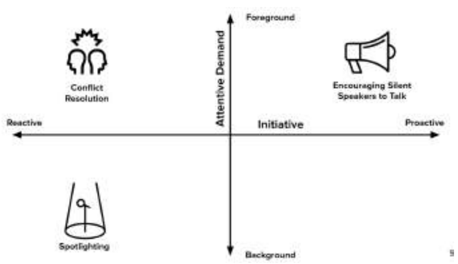

Figure 4: RoboPal Implicit Interaction Framework

DESIGN

A. Behavioral

As highlighted above, RoboPal serves two main functions: to support local conferencing and support video conferencing. Local conferencing involves spotlighting individuals who are speaking, promoting turn taking and equal participation among group members, resolving conflicts as they occur, and discouraging off-topic conversation. In supporting video conferencing, we wanted Robopal to specifically facilitate more face-to-face interactions and also make use of spotlighting to transition between centers of attention. In order to summarize these goals into a more understandable format that we could design around, we made an implicit interaction graph [3]. This graph is based on two axes: the level of attention demanded and the balance of initiative between the user and the system. Robopal's framework is shown in Figure 4. As seen in the framework, RoboPal inhabits several spaces on the graph as it serves several different functions. In its most common use case, it acts as a reactive abstraction automation device in the background as it spotlights speakers. However, in the presence of conflict or as RoboPal takes initiative to encourage equal participation, it draws attention to itself and becomes a commanding force in the interaction.

Following the clarification of our goals for RoboPal, we began to explore many of the behaviors that RoboPal would execute. To do this, we utilized a technique called Wizard of Oz (WoZ). WoZ is a form of role play in which one of our team members act out the movements of the lamp in recreated

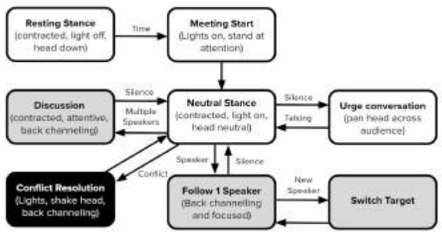

Figure 5: RoboPal State Diagram

Figure 6: Initial WoZ Setup

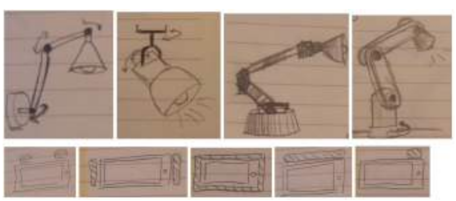

Figure 7: Initial Prototyping Sketches

situations from our storyboards. Other members act as participants and are able to pretend to be in the moment to provide feedback and feelings about their thoughts on the movements. As seen in Figure 5, our first WoZ experiments involved simply zip-tying a stick to a pre-existing lamp we had in our lab. Though simple, this was a very effective way to evaluate movements. After several WoZ sessions, we determined that the robot's head movements were very effective forms of both back-channelling and drawing attention as they closely resembled many human movements such as nodding and shaking. Furthermore, we noticed that vertical movements such as standing up or swaying down was a good way to draw attention from human participants. Our last observation about movements was the importance of the speed and smoothness of the movements. This was easy to do while we were WoZ'ing but we knew that this would pose a difficult challenge with our servo motors.

From our WoZ sessions, we defined several key movements. First, we defined a resting stance and a neutral stance. The resting stance would be a contracted state in which the head lays close to the base, prompting minimal attention from users. As the meeting comes to a start, RoboPal can slowly stand up, turn on its lights, and pan around the room to draw attention. After this, it sits in its neutral stance in which it is slightly bent forward at approximately 90 degrees. As participants begin to discuss, RoboPal will start to follow the main speaker. If a new user starts speaking, the robot can shift its gaze and spotlight the new speaker. If multiple speakers are having a productive conversation, RoboPal can shift into its rest stance and zoom out with the camera to capture both ends of the conversation, fading more into the background as they speak. If that conversation turns into an argument or goes off track, RoboPal can stand at attention and flash its lights or shake its head to draw attention away from the conflict. Lastly, if participants are silent, RoboPal can pan around the room with curiosity to urge conversation. These behaviors formed the foundation of our future testing and served as a map to see how we would WoZ our robot with actual participants. To make this much easier to visualize and understand, we compiled these behavior into a state diagram as seen in Figure 5.

B. Mechanical

As we designed the physical appearance of RoboPal, we went through many waves of brainstorming and prototyping. Initial sketches compared mounting methods (ceiling vs wall vs table-top), arm design, motor actuation (direct vs indirect drive), and the head design. Some of these initial sketches can be shown in Figure 7. This ideation phase was helpful to us because though our chosen design was one of our first visions for RoboPal, looking at alternatives made us sure that this was the design we wanted.

Robopal's ultimate chosen appearance was heavily inspired by the lamp Luxo Sr. from Pixar and the final body is shown in Figure **??**. This involved two longer thin arms and very minimalistic joints. The base was chosen to be a large V shape for maximum stability and to resemble many base designs

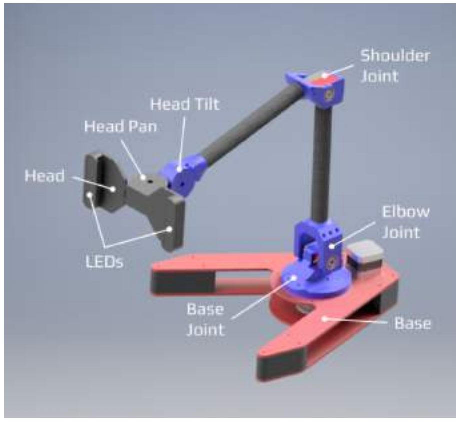

Figure 8: RoboPal Assembly Diagram

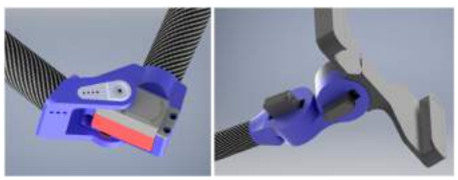

Figure 9: RoboPal's Elbow Joint (left) and Head Assembly (right)

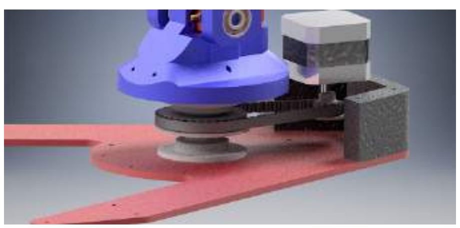

Figure 10: RoboPal's Base Joint Drivetrain

for lamps. The head design was also very important as it is the centerpiece of attention for RoboPal. Through many iterations, our final design was decided to be having two vertical strips of LEDs on the left and right of the phone. This resembled eyes, evoking a subtly anthropomorphic appearance. This was a fine line to walk as we wanted to keep an aspect of looking non-human as well. We wanted users to be able to talk to RoboPal as a participant in local conferences, but to not have the face steal attention from the video participant in remote mode. As mentioned in the previous works, this anti-anthropomorphism discourages users from building a preconceived mental model of its capabilities but would still retain the ability to express humanlike motions with its thin frame "body" and lamp "face."

At the core of RoboPal are its 12" carbon fiber tubes to make up its long and thin body. These tubes were chosen because they are strong, hollow, and light. The joints, base, and head were all manufactured with laser cutting and 3D printing, requiring some raw materials (acrylic sheets and 3D printing filament). We chose to use Adafruit NeoPixels as our lights due to their flexibility and adjustable brightness. This dimming of the lights added another avenue of expression aside from movement.

For actuation, we aimed to implement 5 degrees of freedom to perform critical behaviors and movements, as shown in Figure 8. These DoFs include rotation at the base, extension at elbow and shoulder, and pan/tilt for the head. A NEMA 17 stepper motor drove the base of RoboPal, two high torque (20 kg/cm) servos drove the main joints, and two micro servos were able to drive the head movement. We designed Robopal to be built from 3D printed and laser cut parts. The base plates were laser cut out of $\frac{1}{4}$" acrylic and are separated using 3D printed standoffs. All other components including the joints and head were 3D printed using PLA.

All joints were designed to press fit into the carbon fiber pipes. As shown in Figure 9: the joints integrate their respective servo in the joint to enable for direct drive actuation. The joints were also designed to route the servo wire through the carbon fiber pipes in order to maintain a clean aesthetic. For the elbow and shoulder joints, bearings were used to support the joint on the opposite side of the servo horn in order to maintain structural rigidity. The head assembly was designed to be adjustable in order to accommodate different phone sizes. Half of the head has bars that slide into receiving slots in the other half of the head. The two halves are tensioned together using a spring, which applies enough force to keep the phone secure.

As shown in Figure 10, inside the base is the drive-train for the base joint, which employs a NEMA 17 stepper motor with a belt and pulleys to rotate the main shaft. The main shaft is constrained by bearings mounted to the upper and lower base plates. The shaft extends through the top base plate and fits into the base joint using a square profile. This enables the base joint drive-train to be assembled first before fastening the base joint on top. Because the base joint must rotate the entire arm in any form of extension, we used a 3:1 gear reduction to increase the torque output of the

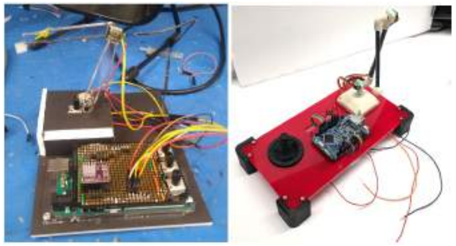

Figure 11: RoboPal controller V1 (left) and RoboPal controller V2 (right)

Figure 12: Testing room layout and view from camera

stepper motor. The belt used was a HTD-5M 270mm, intended for high torque transfer and precision applications. Both pulleys were custom designed and 3D printed to match the HTD 5M tooth profile. The main axle in the base joint has a D-profile that fits into the large pulley in order to transfer torque. The small pulley press fits over the D-profile on the stepper shaft. All hardware used for constraining actuators and general assembly was M3.

After constructing the robot itself, we also had to design an intuitive controller to achieve more organic movements. To do this, we wanted to model the puppet controller used in Tennent's MicBot paper [7]. All movement and light brightness was driven by an NXP FRDM K64 microcontroller dev board. Motion for the base and two arm joints on RoboPal mirrored the potentiometer values for the corresponding joints on a smaller controller. A separate joystick governed the pan and tilt motion of the head and brightness level for the lights. This gives the user complete control over the exact angle of every joint on the arm, and the 'scale model' approach makes it easy to see what geometry will be produced by a given angle. This controller could be hidden effortlessly and provided intuitive controls for conducting WoZ tests with RoboPal. A picture of RoboPal's controller is shown in Figure 11.

TESTING

To evaluate RoboPal's ability to facilitate group problem solving and teamwork performance, 5 groups of 3 to 4 participants each (17 participants total, 12 female, 5 male) were given two problem solving tasks while RoboPal was placed on one end of a table facing them. Participants were comprised of university students between the ages of 17 and 22.

The study was done in Gates 228, a small room with a set of tables and chairs. RoboPal was placed across from the participants on a round table with power cables running back to the power supply on the table at the door. Hidden from the users by the power supplies was the controller we used to WoZ RoboPal. Three video cameras were set up to observe reactions to and interactions with the robot. This overall layout is shown in Figure 12.

Participants were not given information about RoboPal's movements or intent and were blind to the controller. To prompt collaboration and conflict, 2 group members introduced the Island Survival Challenge to participants. In this challenge, the participants are told that they are members of a plane crashing down onto a tropical island. In the plane, there are many objects that may aid them on surviving the island but unfortunately, they can only carry so many of them as they parachute to the ground. They must jump down onto the island in 10 minutes and they must determine the most useful 10 items for survival on the island.

This set up the story and premise of the challenge, as the group had to work together to decide the priority on different items. We printed out pictures of about 30 items including a knife, water bottles, paper, mirror, a pot, a dog, boos, a tent, a flare gun, a saw, duct tape, a lighter, fishing pole, medical kit, and a backpack, to name a few. As the participants talked about their decisions and reasoned through their motives, RoboPal was operating in accordance to the state diagram shown in the behavioral design section.

After the ten minutes were up, we threw in a second portion to the prompt. Participants were told that they each needed to choose a personal item that other members of the group could not utilize. This curveball was to incite further conflict in the group as some members would become selfish and take the more valuable items for themselves. We thought this would be interesting to see because of how different people can make such variable decisions. The differences in mindset and level of confrontation would offer a lot of opportunities for RoboPal to act as a mediator in the interaction.

RESULTS AND CONCLUSION

After conducting our testing, we were able to review the video footage for observations and also get direct feedback from participants through a survey and verbal comments. As a general overview, Table 1 shows the participants of each group, how long their session took, a general level of conflict among participants, how active RoboPal was, and the extent that participants were able to collaborate. Robot activity defines the control of the robot. In half-level activity, we disconnected power from the large torque servos on the arm's joints, as we discovered that their noise and jittery movements could cause distress in situations where the robot was acting more in the background. We would still enable these joints in times where RoboPal needed to take initiative and extend these joints, especially when there were higher levels of conflict and lower levels of collaboration.

Table 1: Summary of Test Results

Group #	# Participants	Duration (min)	Conflict Level	Robot Activity	Collaboration Level
1	3	22	Low	Full	Good
2	4	18	High	Full	Good
3	3	16	Medium	Half	Good
4	3	14	High	Half	Low
5	4	15	Low	Full	Good

After conducting the test, we asked for general user reactions from the participants. In general, users had very mixed feelings regarding RoboPal. On the positive side, users said that RoboPal made them feel more heard, that it felt like it was listening to them talk, and that they were really impressed with the anthropomorphic head movements. However, other comments mentioned the loudness and jittery movements of the large steppers, feeling stressed out and judged by RoboPal, and that it was a little too easy to ignore throughout the activity.

From our survey results, we found that many of our participants found the robot to be "distracting" or "stressful," indicating that they believed their performance was made worse by the robot's presence, even though many of the groups actually showed increased levels of back-channeling and even turn-taking when using Robopal. From this, we conclude that Robopal has positive effects on group dynamics, even when individuals felt negatively about it. We also discovered how critical smooth, quiet motion is in facilitating human robot interaction. Problems we had with the mechanical design of Robopal, including its large form factor and lack of sufficient servo torque, resulted in really loud, jerky movement which was why so many participants were stressed by the robot, and were scared to interact with it. These problems also prevented us from doing testing with Robopal's video conferencing mode. In the future, we can refine our prototype to be smoother and quieter and retest to see how the interaction has changed.

Despite these findings, we also saw a lot of promising results from testing RoboPal. One movement that consistently received positive feedback was the backchannelling. This was done with our most subtle movements of the head, either a nod to affirm, a head tilt to symbolize curiosity, or a shake to draw attention. This is an area that we hope to eventually put additional effort into defining more concise and distinct back channelling movements. Additionally, many users were also able to identify the purpose of RoboPal very accurately, showing that its purpose is intuitive to the users. Many people saw the value in such a device and with better execution in future iteration, we believe there is great potential for RoboPal.

There were also a couple instances where an individual would have especially interesting foreground interactions with Robopal, including some individuals asking Robopal if it was okay, and one individual (M, age 20) who barked at Robopal and engaged with Robopal animalistically. We are unsure of what these behaviors mean, and would like to dive deeper into these particular types of interaction in the future.

Regardless, through the design and testing of RoboPal, we proved that there is a demand for presence in improving these small conference situations. It is crucial that we continue related work as robots become more and more intertwined with our daily lives, as human robot interaction grows.

REFERENCES

[1] Julia Fink. 2012. Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. In *Social Robotics*, Shuzhi Sam Ge, Oussama Khatib, John-John Cabibihan, Reid Simmons, and Mary-Anne Williams (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 199–208.

[2] Guy Hoffman, Oren Zuckerman, Gilad Hirschberger, Michal Luria, and Tal Shani Sherman. 2015. Design and Evaluation of a Peripheral Robotic Conversation Companion. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*. ACM, New York, NY, USA, 3–10. https://doi.org/10.1145/2696454.2696495

[3] Wendy Ju. 2015. *The Design of Implicit Interactions* (1st ed.). Morgan & Claypool Publishers.

[4] Ajung Moon, Daniel Troniak, Brian Gleeson, Matthew Pan, Minhua Zheng, Benjamin Aaron Blumer, Karon Maclean, and Elizabeth Croft. 2014. Meet me where i'm gazing: How shared attention gaze affects human-robot handover timing. *ACM/IEEE International Conference on Human-Robot Interaction*, 334–341. https://doi.org/10.1145/2559636.2559656

[5] Bilge Mutlu, Takayuki Kanda, Jodi Forlizzi, Jessica Hodgins, and Hiroshi Ishiguro. 2012. Conversational gaze mechanisms for humanlike robots. *ACM Transactions on Interactive Intelligent Systems* 1 (01 2012), 12. https://doi.org/10.1145/2070719.2070725

[6] Julia Rozovsky. [n.d.]. The five keys to a successful Google team. https://rework.withgoogle.com/blog/five-keys-to-a-successful-google-team/

[7] Hamish Tennent, Solace Shen, and Malte Jung. 2019. Micbot: A Peripheral Robotic Object to Shape Conversational Dynamics and Team Performance. (01 2019). https://doi.org/10.1109/HRI.2019.8673013

[8] Anita Williams Woolley, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. 2010. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* 330, 6004 (2010), 686–688. https://doi.org/10.1126/science.1193147 arXiv:https://science.sciencemag.org/content/330/6004/686.full.pdf