2024

Evolution of Emotive Behaviors in Simulated Agents

Freyja Feeney
University of Vermont

Evolution of Emotive Behaviors in Simulated Agents

Freyja Feeney
Patrick Leahy Honors College
College of Engineering and Mathematical Sciences
Dr. Joshua Bongard, Amanda Bertschinger

April, 2024

Contents

Abstract

As our lives become increasingly intertwined with technology, it becomes imperative to study human-robot interaction (HRI) through an emotional lens. Humans relating to robots on an emotional level could lead to enhanced robot-human collaboration and broader acceptance of technology. In the future, robots that perform dangerous rescues, care for the elderly and act as therapists will all benefit from the ability to display interpretable emotional behaviors. The purpose of this project was to develop simulated robotic behaviors that are expressive and representative of certain human emotions using evolutionary robotics. The result is an emotive behavior evolved from a random behavior that people can then recognize as a specific emotion. The robot designed for this project was quadrupedal and doglike. It has a 47-neuron neural network with 16 sensor neurons, 15 motor neurons, and 16 hidden neurons. Through anthropomorphism, human subjects were able to view an evolved robotic behavior from the robot, and come to a significant consensus on what emotion the behavior might represent. These results indicate that people are capable of agreeing on emotional robotic behaviors. With the advancing integration of robots into society, the capacity of humans to universally interpret robotic behaviors becomes increasingly paramount in tandem. This research proves for the first time that it is possible to evolve a robotic behavior that people can collectively recognize as expressing a certain emotion.

Definitions and Acronyms

Evolutionary Algorithm: Algorithms modeled after natural evolution used to solve optimization and search problems

Morphology: The physical shape/build of the robot's body

Link: The cubes that make up the body of the robot

Joint: The edge where the links are connected in the robot

Controller: The robot brain evolved by the evolutionary algorithm

Fitness Function: Equation that the evolutionary algorithm uses to evolve and evaluate the success of the behavior

Embodied Cognition: The idea that intelligence comes also from interactions with the environment through a physical body

HCI: Human-Computer Interaction

PHC: Parallel Hill Climber

HCR: Human-Chatbot Relationships

HRI: Human-Robot Interaction

1. Introduction

Human-computer interaction (HCI) is a massive field that has rapidly advanced in recent years. However, the emotional aspect of HCI remains relatively unexplored in comparison. There has been some work done on evolving emotions in robots using evolutionary robotics, specifically fear. “Basic emotions and adaptation. A computational and evolutionary model” describes the attempt to evolve fear responses in agents to respond and react appropriately to dangerous situations (Pacella, Ponticorvo, Gigliotta., & Miglino, 2017). However, rather than evolve or wait for the emergence of organic emotions in a robot, this study looks at the emotions people attribute to robot behaviors through anthropomorphism. “Anthropomorphism and the Social Robot” defines anthropomorphism as the tendency for humans to attribute emotions to inanimate objects and animals to better understand their behavior (Duffy, 2003). There have been studies such as “Anthropomorphism of computers: Is it mindful or mindless? Computers in Human Behavior” that reveal that humans do indeed anthropomorphize and empathize with robots. And also that this behavior can happen unconsciously, or mindlessly (Kim, Sundar, 2012). The reality is that while machines do not feel human emotions, if they behave in a certain way, and look a certain way, the human brain will do all of the work. Anthropomorphism can be used in HCI to facilitate more comfortable, understandable interactions. Duffy describes innovative possibilities for the social robot, including the use of biosensors in humanoid robots that gauge stress and respond accordingly. For example, Pepper was a humanoid robot capable of reading human emotions. However, production was stopped in 2020 due to a lack of demand (Crowe, 2021). Rather than designing a humanoid robot that can recognize human emotions, this project focuses on utilizing the emotions people subconsciously project onto robots to begin with. It will explore whether or not an evolutionary algorithm can evolve an emotive robotic behavior that a crowd can interpret.

1.1 Evolutionary Robotics

Evolutionary robotics is a sub-discipline of artificial intelligence that focuses on the design of robot morphology using the principles of natural evolution. Evolutionary algorithms are employed to autonomously develop robot designs, morphology, and control systems, to complete a desired goal, ranging from locomotion to adaptability. Here, we limit our evolution to the robot control systems in order to explore the behaviors these control systems create. In looking at ways to develop emotive robot behaviors, evolutionary robotics provides many advantages over traditional engineering. Evolutionary robotics encourages, through its iterative nature, the exploration of a wide range of behaviors that could better capture the nuances and intricacies of emotional body language.

1.2 Designing the Morphology

In beginning this research, the first question that arose was "What does it mean to be happy?" In the context of physical behaviors and body movements, what does it mean to feel and portray a certain emotion? The importance of the head and neck became evident immediately. The position of the head on almost any morphology, animal-like or humanoid, was a strong indicator of what emotion the entity appeared to be feeling. Because of this, it was decided that the robot robot design must include a head and a neck. Other attributes, like gait and distance of the body from the ground, were similarly effective in portraying emotion. "Anthropomorphism of Robots: Study of Appearance and Agency JMIR Hum Factors" describes the idea that humans are less likely to anthropomorphize humanoid robots because people have a high sensitivity to other human appearances and can notice, and be made uncomfortable by, differences even at an infinitesimal level (Crowell et al. 2019). To avoid this effect, the robot's morphology was designed to resemble a dog, with four legs, a torso, a neck, and a head.

Figure 1.

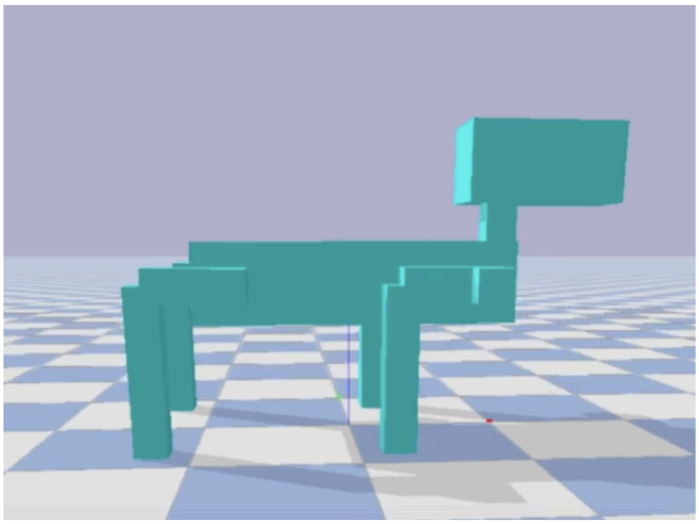


Figure 1. Image of robot morphology

1.3 Anthropomorphism of Robots

It is well known that humans tend to attribute human emotions and feelings to non-human entities, like animals and pets; this is called anthropomorphism. When considering ways to approach and possibly cultivate an emotional relationship between humans and computers, it made sense to begin with an investigation of anthropomorphism. Rather than attempt to design a robot that can sense human emotional queues, register them, interpret them, and then respond accordingly, this study utilizes the innate human ability to feel emotions and attribute emotions to non feeling entities. By studying what behaviors people have emotional responses to, there is no need for the robot to attempt to understand the complexities of human emotion. Rather, it is equipped with behaviors that are known to be emotionally interpretable to users. By comparison, this is a more time and cost-effective way to integrate emotions into an interaction. These behaviors would be useful in many real-world scenarios, for example, robots that work in healthcare. If a robot is equipped with a set of behaviors that are known to elicit happy emotional responses from children, they are capable of comforting patients in an emotional way. Robots that perform rescues could be equipped with behaviors that appear trustworthy, and be able to convince a scared person to follow them out of a burning building. Therapist robots could have body language that people see as safe and familiar. The goal of this project is to develop these interpretable, emotional behaviors using evolutionary computation. There is some previous work on the anthropomorphization of computers, and it has been proven that people can empathize with robots. This is demonstrated, for example, in the extreme by intimate human-chatbot relationships (HCR) seen as often in reality as in science fiction (Skjuve, Følstad, Fostervold, & Brandtzaeg, 2022).

2. Methods

The evolutionary algorithm was coded using Python 3.10 and Visual Studio Code. The robot and environment were simulated using Pybullet, a Python-based physics engine. The Pyrosim package was used to more efficiently send and receive information from Pybullet. This project used the Parallel Hill Climber (PHC) algorithm. The PHC is a variant of the Hill Climber, which is a local search algorithm used to find the best solution within a given scope. The search begins with a randomly generated solution from which neighboring solutions are generated and evaluated against an equation. These comparisons happen in parallel, multiple are evaluated simultaneously, hence the parallel in Parallel Hill Climber. Chrome remote desktop was used to communicate between the main laptop and a Windows PC used for Twitch streaming.

The consensus on the behaviors was gathered first via surveys, with the process most closely resembling a single-blind experiment. The participants were unaware of the goals and intentions of the experiment. To test if people could recognize emotion in our simple, faceless, quadrupedal robot, we first generated a series of unevolved behaviors where the objective was maximizing the x distance. These attempts yielded a couple of random controllers that happened to be emotive behaviors according to us. The initial emotions chosen were happy, sad, angry, frightened, and pained. However, to adhere to certain human subject testing ethicality practices, a few of the more negative emotions were removed. The three behaviors that drew the most consensus were happy, sad, and lazy. The happy behavior was categorized by a raised head and quick, jumping steps. The sad behavior was categorized as a lowered head and a body hovering closer to the ground, and the lazy behavior was lying down and not moving. The second and third round of surveys sent out were to make sure the engineered fitness functions were replicating the behaviors from the first survey, and were still able to gather consensus. We wanted to ensure we were properly evolving for the behaviors in the videos, and sent out surveys to prove that we had properly translated the behaviors into fitness functions that would effectively represent the emotions. By just observing the random controllers and writing them into functions ourselves, we weren't able to ensure that they are close enough to the original behaviors to end up with the same degree of consensus. So, we designed the non-random controllers and sent out another round of surveys.

All surveys displayed clips of the robotic behavior to the viewer, and multiple-choice questions asking which emotion the robot appeared to be feeling. Each question in the survey consisted of a video of a behavior and a set of multiple-choice answers. The questions were worded carefully to best avoid bias and to adhere to Institutional Review Board (IRB) guidelines on human subject testing. The survey questions focused on asking what the viewers believed the robots were feeling, not how the robots made the viewers feel. The answer choices also contained emotions that weren't shown in any of the videos, to remove some of the bias that might come from the process of elimination. The forms were sent to friends and family with no other context than please watch the video and respond to the questions. The majority of the audience had no prior experience in robotics or engineering.

3. Results

3.1 Preliminary Surveys

The first survey was sent out to friends and family. The figures below show the results for excited, sad, and happy. First survey responses for sad, happy, and excited, all show consensus or majority. With 16 responses, sad had a 93.8% consensus, and excited had 68.8%. Of the 15 responses for happy, there was a 53.5% consensus.

Figure 2a

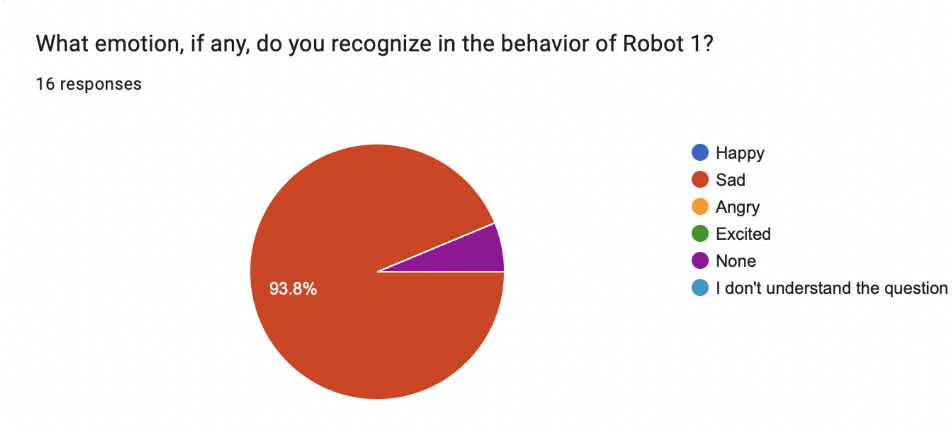


Figure 2b

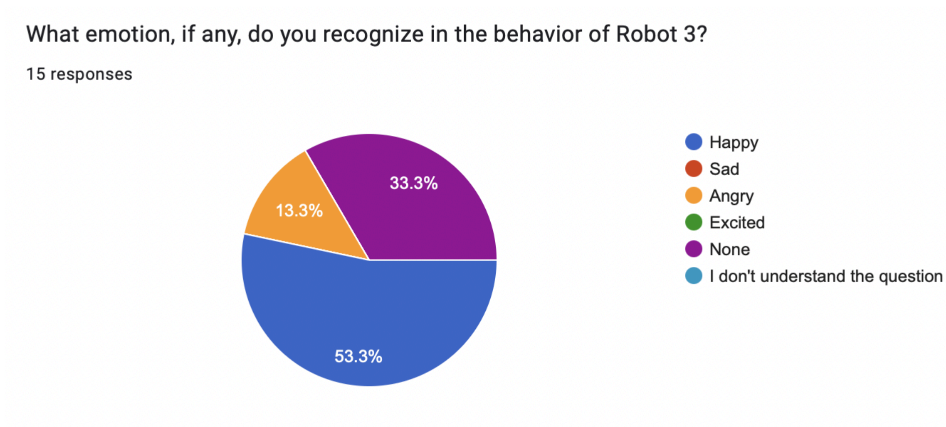
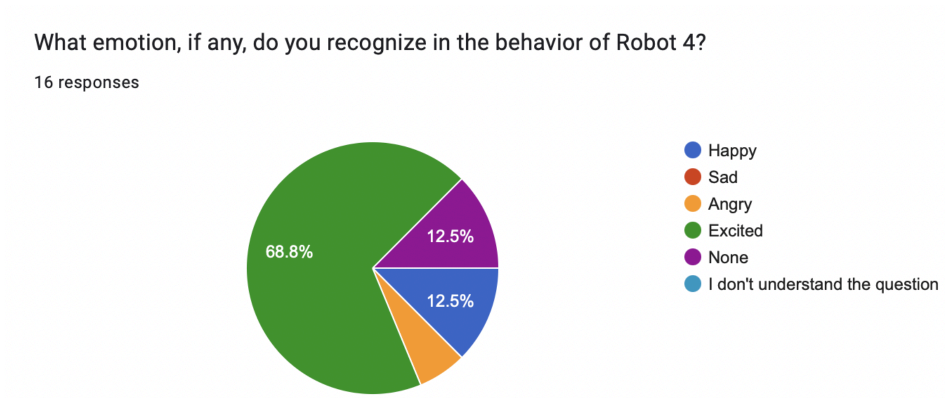


Figure 2c



The second survey collected 22 responses and yielded results for another set of emotions, with the relevant ones being sad at 63.6% and happy at 81.8%.

Figure 3a

Second survey results show consensus for sad and happy.

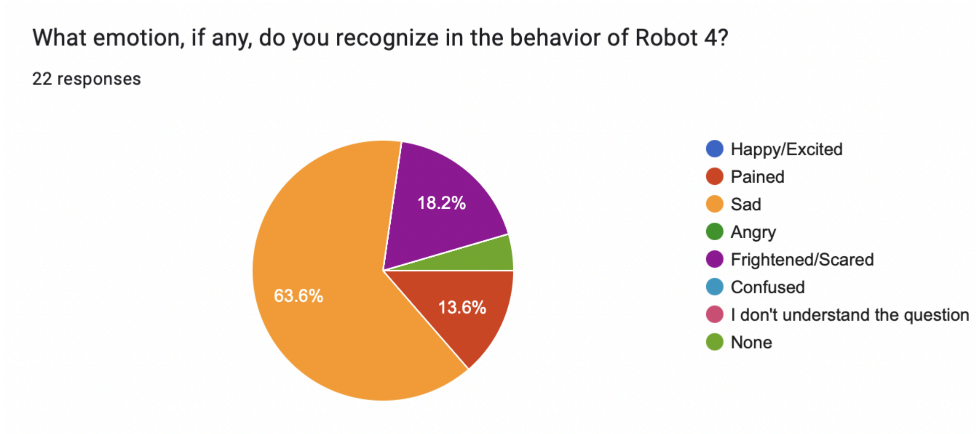
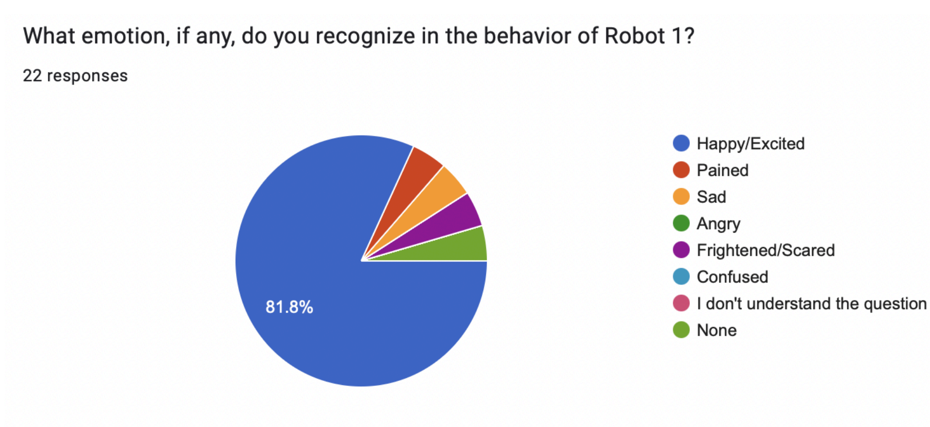


Figure 3b



After the two preliminary surveys, it was concluded that people can recognize and agree on emotional behaviors in our robot. The behaviors with the most consensus were then analyzed to design a fitness function that could replicate the behavior from the static videos. For example, if the sad behavior had a lowered head, an engineered fitness function would select a head closer to the ground. They were engineered so that the behaviors from the videos could be recreated and possibly placed into different robot morphologies, and eventually, to be used in interactive evolution.

The third survey had nine responses for the final emotions: happy, sad, and bored/lazy. The fitness functions for the three emotions were designed based on the behaviors noted from the static random videos. For example, if the sad behavior had a hanging head, the fitness function for the head selected for the position of the head link to have a negative z-position, or to be closed to the ground.

Figure 4a

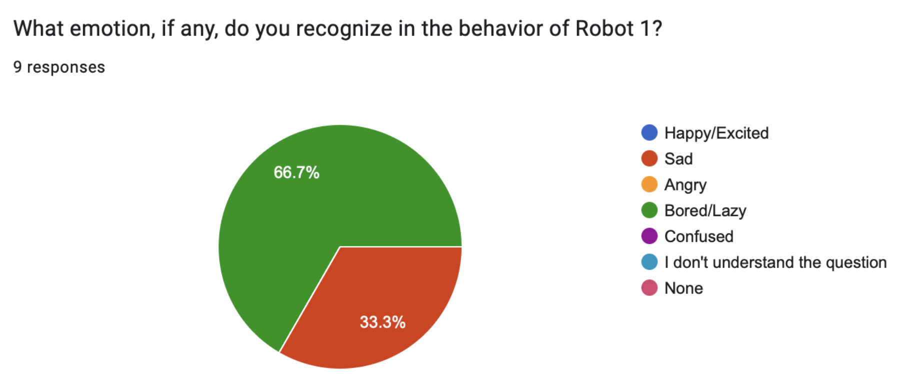


Figure 4b

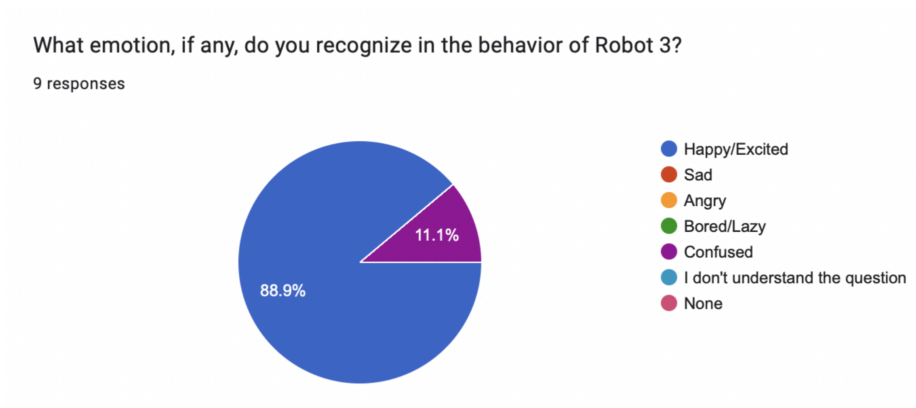
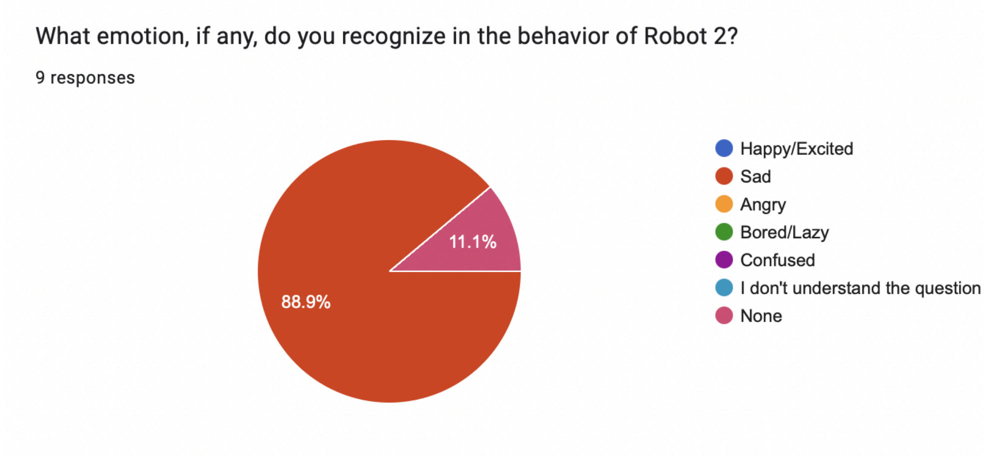


Figure 4c



The three fitness functions were as follows:

1. Happy:

   *total = (self.xPosition * 0.5) + (self.z_position_of_head * 0.9) + (self.zPosition * 0.5) + (joint_chance * 0.05) + (-force_change * 0.3)*

2. Sad:

   *total = (self.xPosition * 0.3) + (-self.z_position_of_head * 0.9) + (self.zPosition * 0.3) + (-joint_change * 0.05) + (-force_change * 0.7)*

3. Lazy/Bored:

   *total = (-self.xPosition * 0.01) + (-self.z_position_of_head * 0.9) + (-self.zPosition * 0.5) + (-joint_change * 0.05) + (-force_change * 0.9)*

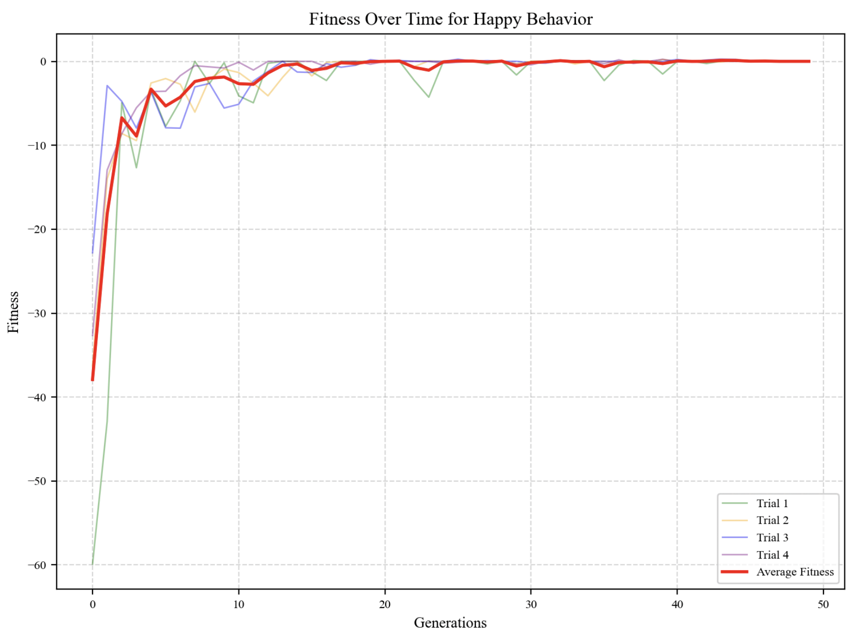
- **root link:** first link in the robot from which the other links extend, in this robot it is the "Torso" link
- **self.xPosition:** is the position of the root link of the robot on the x-axis, in a three-dimensional Cartesian coordinate system, negative to the left and positive to the right in the simulation
- **self.z_position_of_head:** the absolute position of the head link, with positive being higher up and negative being lower to the ground
- **self.zPosition:** the up and down position of the root link of the robot, with positive being higher off the ground and negative being closer
- **joint_change:** the change in the angle of the joint between iterations from its expected value

- **force_change:** the force the motors exert, how strong the robot is

These functions are a weighted combination of parameters, joint positions, and changes in force that represent behaviors that express the three different emotions. Over time, the behaviors evolved, as shown in Figures 1-3 below.

Figure 5

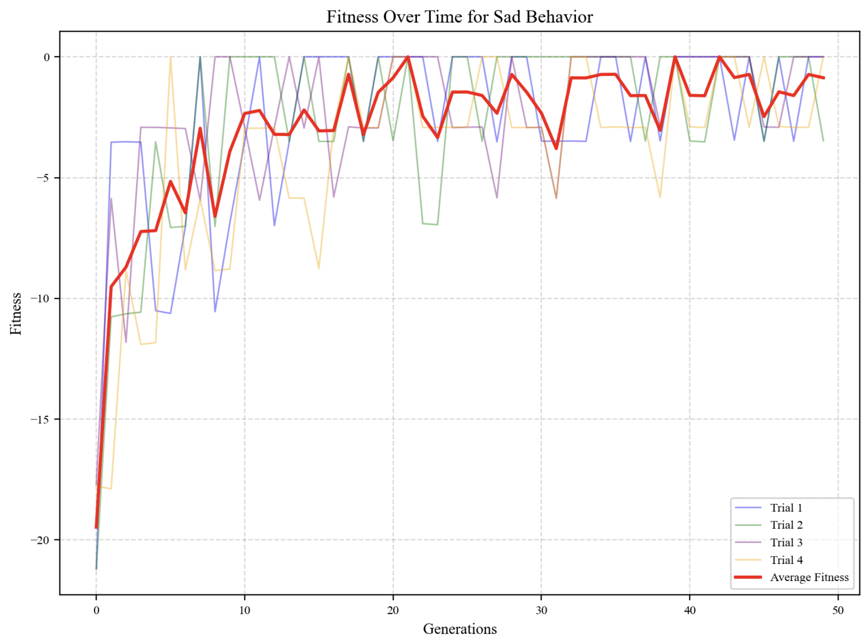
Happy Behavior



Note. 50 generations, average of the of population 10 per generation, 4 trials with average in red

Figure 6

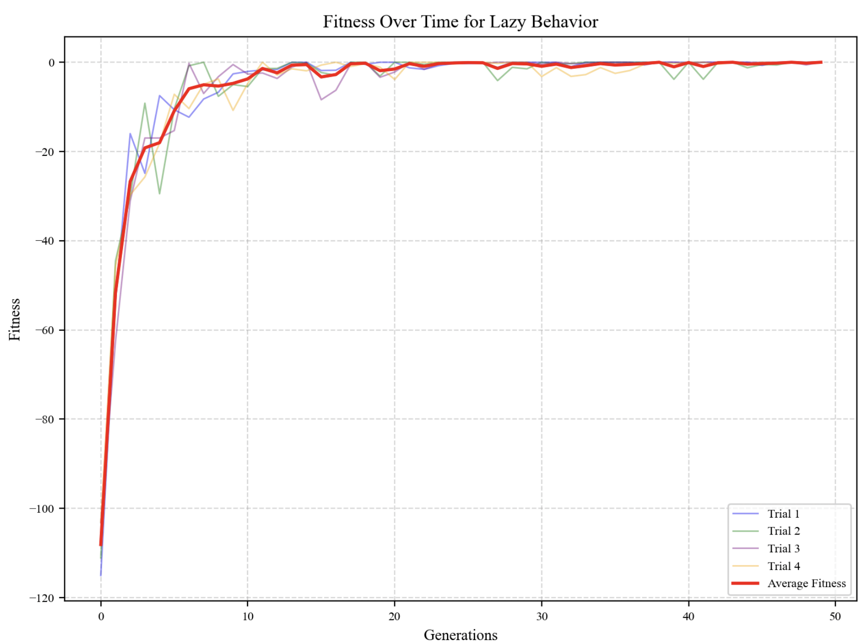
Sad Behavior



Note. 50 generations, population 10, 4 trials with average in red

Figure 7

Lazy Behavior

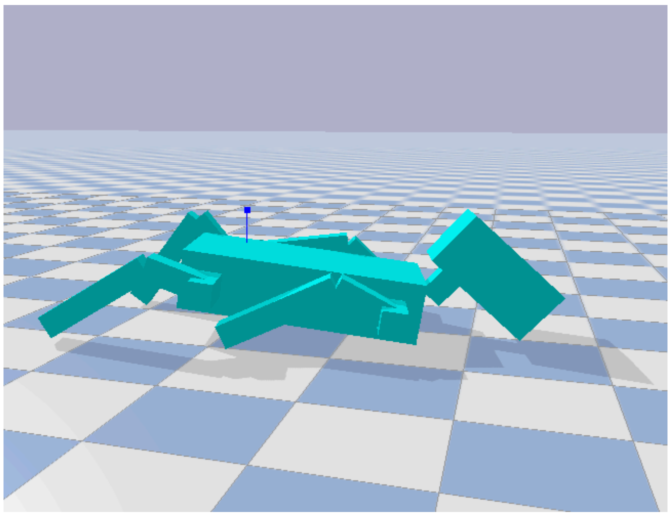


Note. 50 generations, population 10, 4 trials with average in red

Over time, the robot evolves from a random behavior to a behavior that successfully portrays the target emotion. The behaviors in the figures above evolved for 50 generations with a population of 10 for four trials. The algorithm used was PHC, and the average fitness across the 10 individuals in the population was stored for each generation.

Figure 8

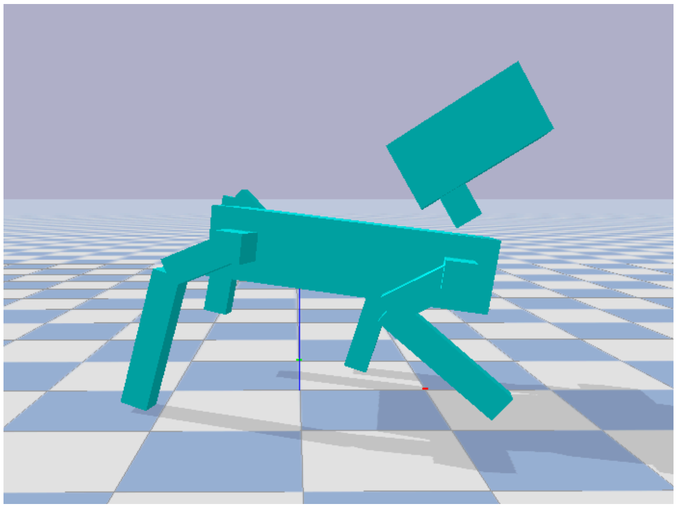
Screen Capture of Lazy Robot



Note. Generations: 20, Population: 5

Figure 9

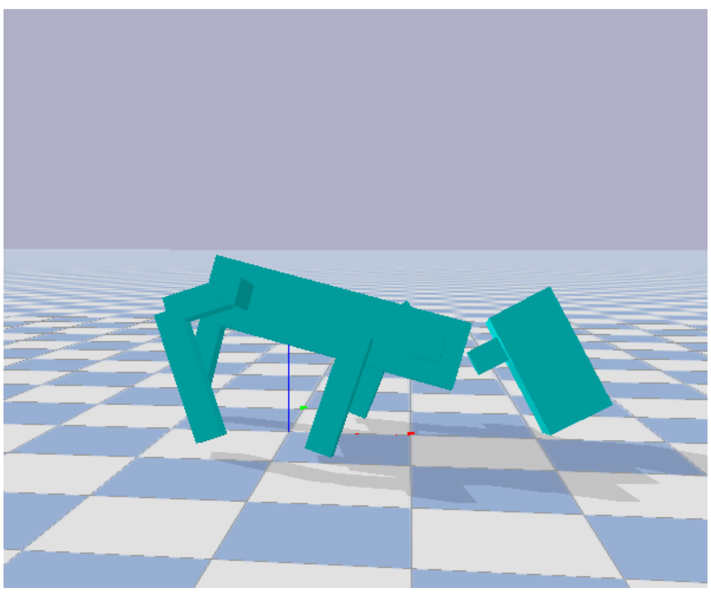
Screen Capture of Happy Robot



Note. Generations: 20, Population: 5

Figure 10

Screen Capture of Sad Robot



Note. Generations: 50, Population: 10

4. Discussion

Summary

It was proven that it is possible both to evolve emotional behaviors in robots and for crowds to come to a consensus on what emotions those behaviors portray, visible in Figures 2a through 4c.

It was shown that it is possible to study emotive behaviors and translate them into a fitness function that can be used to evolve controllers. The surveys all showed a higher than majority agreement on the emotions of interest and proved that it is possible to come to a consensus on the implications and meaning of robot behaviors in an emotional context. The amount of consensus varied, but was up to as high as 93.8% (Figure 2a). It is possible to evolve, as seen in Figures 5 through 7, a robotic behavior that is interpretable to humans.

Conclusions

It is a significant discovery that a crowd can come to an agreement on something as subjective as an evolved emotion in something as objective as a robot. This is essential for the integration of any sort of emotional robot into society. The context and implications of emotive robotic behaviors must be understood by the general public in the same way, in order for the robot to do its job effectively and safely. Misunderstandings of robotic behaviors would not only damage trust between humans and robots but pose as potentially dangerous as well. It is important to study what affects consensus in the interpretation of robotic behaviors, and what robotic behaviors might be universally representative of a certain emotion. This research proved that it is possible to evolve an emotive behavior that can draw consensus from a crowd. This crowd, as well, should consist of a wide variety of people in order to be representative of the general public. A contained group of engineers that design and model robotic behaviors would not be representative of the way the public may interpret robotic behaviors.

This research addresses a gap in human-computer interaction today by bringing in emotional interactions between humans and robots, and exploring how to design and evolve emotionally expressive behaviors that humans can recognize. In the design of robots in the future, it is important to not only look at the feasibility of the behavior from a technical standpoint but to also make sure it is interpretable and understandable to humans. Evolving these behaviors is one way to do this. Crowd consensus on emotional behaviors holds promise for application across diverse domains such as healthcare, entertainment, and elderly care. The behaviors can be used to build trust between humans and robots and to increase the effectiveness of robot-human collaboration. And, because the behaviors are interpretable and understandable to crowds, the robots could be more smoothly integrated into our everyday lives. Not only because they make sense to people, but also because they will fit into our legal systems in a more natural way than a robot with foreign and unintelligible behaviors developed by a single team of engineers.

5. Future Work

5.1 Interactive Evolution

Eventually, the idea of drawing crowd consensus on behaviors evolved into using crowd-controlled feedback to develop the behaviors themselves. If a crowd were able to collectively shape an emotive behavior, it would be universally interpretable and a much better representation of how the general public might perceive emotive behaviors. Utilizing interactive evolution would open up the development of these behaviors, removing the need for an intermediary engineer, and facilitating the communication between the evolutionary algorithm and natural language feedback from people. This is much more representative of the general public and opens up the possibility of accounting for variances in perception due to differences in race, religion, ideologies, etc. This interactive component would be done using the streaming platform Twitch, and a chatbot that would take responses from the chat and feed them into the evolutionary algorithm. For example, an evolving emotive behavior would be displayed to the viewers of the stream as an MP4, and the viewers would be prompted for the success of the behavior using some scale. The algorithm would then adjust based on the approval of the users in the chat, and evolve accordingly until a behavior with a significant consensus in the chat is created.

This would be working off of a previous project from the Morphology, Evolution, and Cognition Lab, called Twitch Plays Robotics (Bongard, Anetsberger, 2016). This project's objective was to use crowds to try to teach robots language. Similarly, the chat was used to take in responses from the crowd, and those responses directed the evolution of the robots. A chatbot that connects to Twitch and plays a video of an emotive behavior on loop was created as well. The bot collects poll data from the chat, regarding which emotion the robot seems to be portraying, and outputs the consensus as a percentage. By using a public channel on Twitch, the crowd was expanded from a team of programmers and their friends and family to anyone on the internet with a Twitch account. This would effectively eliminate the bias that comes from being invested in the success of the experiment, allowing for a more objective and accurate representation of how the general public might react. Preliminary results yielded a consensus for the sad behavior. It was difficult to generate traffic to the Twitch channel as it was new and had no following, but it offers the potential for millions of unbiased users to view and interact with the experiment. The stream was live for 186 hours and had 44 unique viewers and 7 unique chatters. While some of the chatters were most likely bots, the stream attracted 44 viewers with relatively no advertising, which is already double the original sample size of friends and family. The link to the stream was put up on social media near the end of this experiment and is likely to attract more traffic.

Figure 11

Screen Capture from Twitch Stream

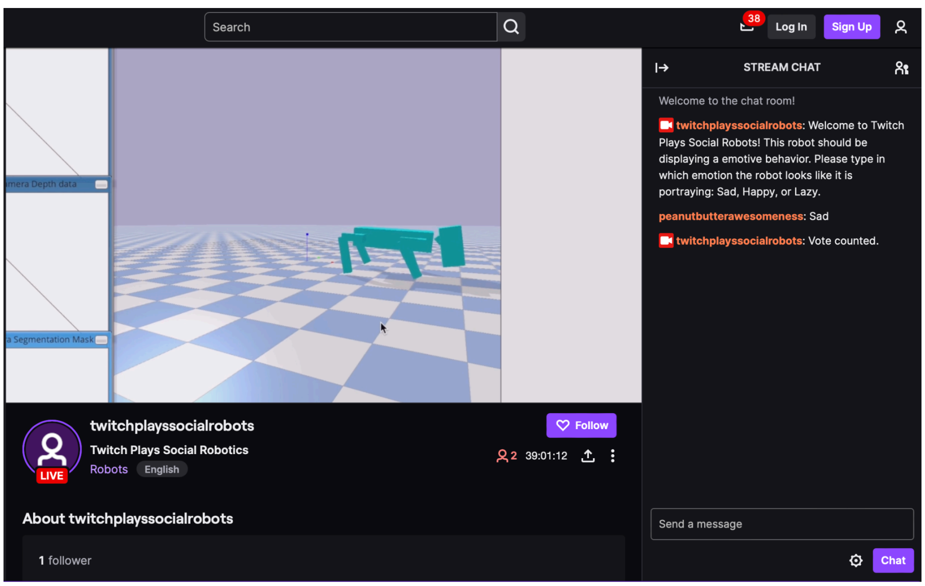


Figure 12

Output From Chatbot Polling

```
sad count: 2
consensus: 100.0%
Received message: aw he looks sad
sad count: 3
consensus: 100.0%
Received message: Too slow
sad count: 3
consensus: 100.0%
```

Preliminary interactive features were implemented into this project, locally first. In the terminal, the user was prompted for which emotion they would like to evolve for. The algorithm would then run and pause every five generations, and prompt for the success on a scale of 1-5 of that behavior. The fitness function was then scaled by that rating, thus taking the approval of the user into account for the success of the controller.

Figure 13

Preliminary Interactive Features

```
pybullet build time: Oct  2 2023 20:25:04
What emotion do you want to see?
        1. Happy
        2. Sad
        3. Lazy
```

```
43.88659811456175
44.0255458180635
43.51197405368376
45.15383212091371
best fitness: 45.15383212091371
Generation 0, What emotion does the robot look like it is portraying?:
```

We began work on the interactive elements, and so far have a Twitch chatbot that polls the chat for consensus. After ten votes, the algorithm is rerun, the behavior re-evolved with consensus scaling the fitness function. Pybullet then outputs an .mp4 file of the resulting behavior. This cycle should continue and the behavior would continue to evolve until, ideally, consensus reaches around 90% (9/10) among the participants. The consensus on the behaviors must be high because the objective of the experiment is to see if it is possible to use crowdsourcing to develop universally intelligible emotive behavior in simulated agents. The ability of people to recognize emotional behaviors and body language is also typically quite high. This study would explore whether or not the same universal rules apply to emotive behavior in robots and non-humanoid entities. The video is then streamed on the Twitch channel, and again the chat is polled for the success of the behavior. This system would be reusable, as it would be a way to develop emotional behaviors using crowds that are sure to be interpretable, without the need for a software engineer or any knowledge of robotics in general.

6. Limitations

The way humans perceive their environments can be shaped and influenced by their past experiences, potentially introducing bias into our surveys. For instance, if an individual has had many negative experiences with dogs, they are more likely to perceive negative or intimidating emotions from a dog-like robot. This should be kept in mind when designing and selecting robot morphologies that resemble animals or other organisms people often interact with. Because our survey included videos of an animal-like robot and focused on anthropomorphization, it is possible that the responses were influenced by this kind of bias. However, we concluded that a robot designed to look like a creature that is familiar, even when unpleasant, is arguably more comfortable to interact with than an unrecognizable or alien morphology.

The three rounds of surveys also had some overlap in terms of people who responded. Both members of the team sent them to friends and family, with no overlap between those two groups. There were likely several respondents who were exposed to all three rounds of surveys, and this familiarity may have influenced the answers to the questions. However, the surveys were spaced several weeks apart, and no further information was provided from one round of surveys to another. In the future, a larger, more diverse pool with more variation between surveys should be used in order to mitigate these biases.

Acknowledgments

Thank you to my advisor, Dr. Josh Bongard, for all of his guidance throughout the course of this project and my time at UVM. I would also like to thank Amanda Bertschinger for being an incredible mentor and big contributor to the direction and success of this research. Thanks to Lisa Dion for being on my committee and supporting me during my undergraduate career as well! And thanks to the Morphology, Evolution, and Cognition lab for immersing me in the world of computational evolution and embodied cognition.

Works Cited

Brian R. Duffy, Anthropomorphism and the social robot, Robotics and Autonomous Systems, Volume 42, Issues 3–4, 2003, Pages 177-190, ISSN 0921-8890, https://doi.org/10.1016/S0921-8890(02)00374-3.

Crowell C, Deska J, Villano M, Zenk J, Roddy Jr J Anthropomorphism of Robots: Study of Appearance and Agency JMIR Hum Factors 2019;6(2):e12629 URL: https://humanfactors.jmir.org/2019/2/e12629 DOI: 10.2196/12629

Crowe, S. (2021, June 28). Report: Softbank stopped production of Pepper robot in 2020. The Robot Report. https://www.therobotreport.com/softbank-stopped-production-of-pepper-robot-in-2020/

Cynthia Breazeal, Emotion and sociable humanoid robots, International Journal of Human-Computer Studies, Volume 59, Issues 1–2, 2003, Pages 119-155, ISSN 1071-5819, https://doi.org/10.1016/S1071-5819(03)00018-1.

E. Tuci, T. Ferrauto, A. Zeschel, G. Massera and S. Nolfi, "An Experiment on Behavior Generalization and the Emergence of Linguistic Compositionality in Evolving Robots," in IEEE Transactions on Autonomous Mental Development, vol. 3, no. 2, pp. 176-189, June 2011, doi: 10.1109/TAMD.2011.2114659.

H. Takagi, "Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation," in Proceedings of the IEEE, vol. 89, no. 9, pp. 1275-1296, Sept. 2001, doi: 10.1109/5.949485.

Josh Bongard, Joey Anetsberger; July 4–6, 2016. "Robots can ground crowd-proposed symbols by forming theories of group mind." Proceedings of the ALIFE 2016, the Fifteenth International Conference on the Synthesis and Simulation of Living Systems. ALIFE 2016, the Fifteenth International Conference on the Synthesis and Simulation of Living Systems . Cancun, Mexico. (pp. pp. 684-691). ASME.

Pacella, D., Ponticorvo, M., Gigliotta, O., & Miglino, O. (2017). Basic emotions and adaptation: A computational and evolutionary model. PLOS ONE, 12(11), e0187463. https://doi.org/10.1371/journal.pone.0187463

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2022, August 3). A longitudinal study of human–chatbot relationships. International Journal of Human-Computer Studies. https://www.sciencedirect.com/science/article/pii/S1071581922001252

Youjeong Kim, S. Shyam Sundar, Anthropomorphism of computers: Is it mindful or mindless?, Computers in Human Behavior, Volume 28, Issue 1, 2012, Pages 241-250, ISSN 0747-5632.

Zhenyu Gu, Ming Xi Tang, John Hamilton Frazer, Capturing aesthetic intention during interactive evolution, Computer-Aided Design, Volume 38, Issue 3, 2006, Pages 224-237, ISSN 0010-4485, https://doi.org/10.1016/j.cad.2005.10.008.