UNIVERSITY OF LOUISVILLE

LOUISVILLE AUTOMATION & ROBOTICS RESEARCH INSTITUTE

KY MULTISCALE
The NSF NNCI Multi-Scale Manufacturing & Nano Integration Node

National Nanotechnology Coordinated Infrastructure

Physiological Signal Analysis During Human-Robot Interaction for Children with Autism

Rohit Narayanan[1], Sai Vanaparthy[2], and Karla Conn Welch[2]

[1]Department of Electrical and Computer Engineering, Princeton University [2]Department of Electrical and Computer Engineering, University of Louisville

Machine Learning & Interaction

M L I

Introduction

Affect-sensitive human-robot interaction is the continuous feedback of information regarding the subject to the robot so that the robot can vary its behavior and respond to changes in the subject's mood and attentiveness.

This can be applied to the treatment of autism by having robots serve as social skills enforcers by stimulating and encouraging conversation, while monitoring the subjects through non-invasive sensors and adapting its program.

We tested the ability to gather physiological data from non-invasive sensors through interactions with a NAO robot (Fig. 1) monitored by an Empatica E4 wristband (Fig. 2).

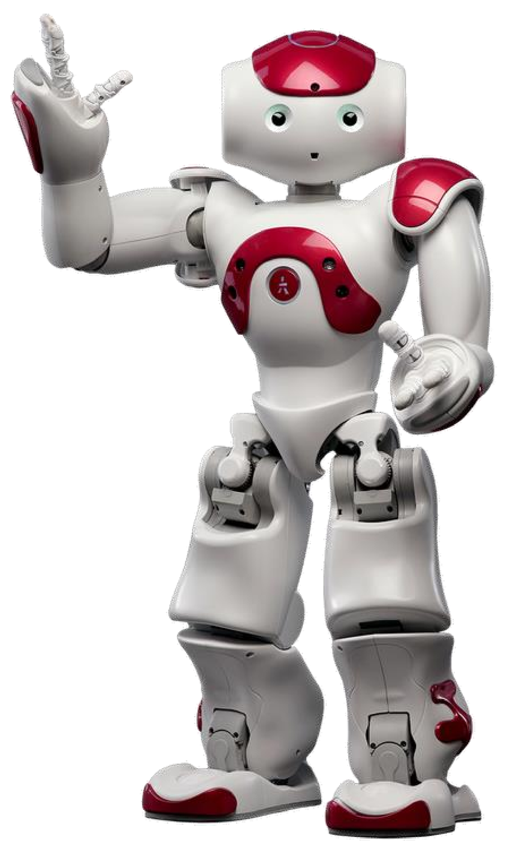

Figure 1. NAO robot

Figure 2. Empatica E4

Experiments

Experiment 1: Fall 2020 Autism Center Study

- Two male children, 11.5 years and 11.25 years old, with autism wore E4s and interacted with the robot (Fig. 3). They practiced short conversations by asking questions to NAO off a prepared list.
- E4 data was collected for 12 sessions of chat time with peers (without a robot) and 10 sessions of practicing conversation skills with a robot.
- Videos of the sessions were coded for 20s blocks representing *inattentive* or *attentive* behavior.

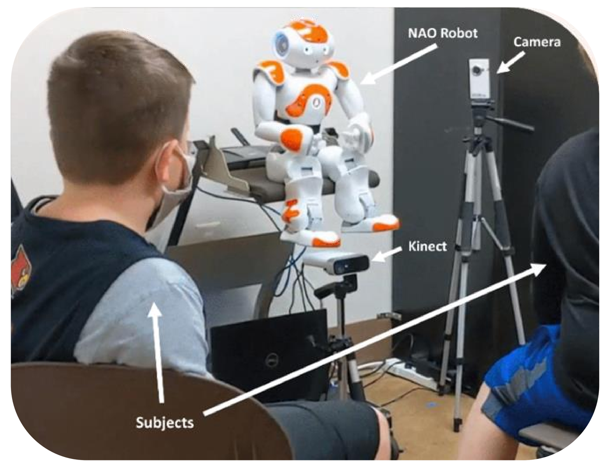

Figure 3. Fall 2020 Autism Center Set-up

Experiment 2: User Acceptance Study

- 17 (5 male) adult subjects (M = 22.1 years, SD = 3.93) with little to no robot experience answered RoSAS and Godspeed surveys after interacting with NAO.
- Initial impressions of four conditions, combining two motions (Jerky vs. Smooth) and two voices (NAO default vs. Justin from Amazon Polly AWS), were tested.

Conditions: *A – Smooth NAO* *B – Smooth Justin*

Conditions: *C – Jerky NAO* *D – Jerky Justin*

- Via Godspeed, Condition B was perceived as *most safe* and *most human-like* while Condition C was rated as *least safe* and *least human-like*, with significant t-test results for safety (p=7.8x10^{-5}) and anthropomorphism (p=8.4x10^{-3}). RoSAS showed Conditions C and D induced discomfort in subjects, supported by ratings on *strange* and *awkward*. Friedman's test also found significance (α<0.05) in both surveys.

Experiment 3: Fall 2021 Autism Center Study

- Two subjects will be placed in a room and seated across from each other, with a robot between them (Fig. 4). They will practice conversations.
- NAO gives prompts (Fig. 5) upon detected silence or one subject dominating the conversation [1].

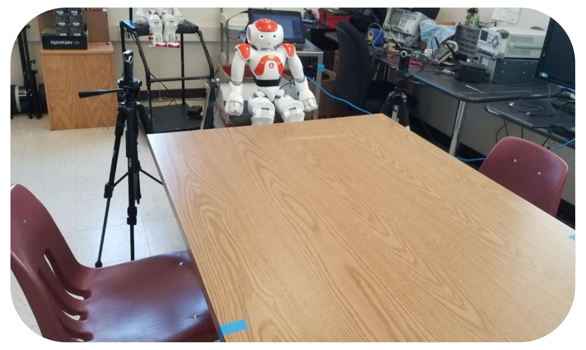

Figure 4. Proposed set-up for study in Fall 2021

Methods

For each experiment, the E4 physiological data is compiled and fed into a machine learning algorithm by the following process:

1. Raw E4 data converted to .csv file with timestamps using a Jupyter function
2. Condition start and end time stamps calculated through manual video analysis
3. Use an IIR Filter to divide tonic and phasic portions of EDA
4. Use a Bandpass Filter on the raw BVP signal
5. Segment into 20 second blocks
6. Calculate 12 features [2]
   a. EDA Tonic Mean, Standard Deviation
   b. EDA Phasic Peak Rate, Max Amplitude, Mean Amplitude
   c. BVP Peak Mean Amplitude, Max Amplitude
   d. HR Mean, Standard Deviation
   e. Skin Temperature Mean, Slope, Standard Deviation
7. Divide features by baseline (lowest deviation 20 second block) [3]
8. Normalize all segments between zero and one
9. E4 segments are matched either to condition (User Acceptance) or to a category given by a trained human coder (Autism Center)
10. If the segments are coded, 30% are re-coded to judge interrater reliability using Cohen's Kappa [4]
11. Data is balanced between different categories using SMOTE oversampling and fed into a machine learning module to calculate results from five separate algorithms
    a) Logistic Regression, Support Vector Machine, Decision Tree, Random Forest, and Gradient Boosting

Results – User Acceptance Study

The F1 score is a weighted average of precision and recall, ranging from zero to one. A statistically significant result would be F1 > 0.5 in all categories. Overall Accuracy represents the total number of correct predictions over total number of predictions.

For experiment 1, results on F1 scores for user-dependent models, which create individual models for each subject, have been calculated. The models for each subject have an accuracy for predicting each behavior (inattentive and attentive) above 0.67.

Experiment 2 tested a user-independent model, which combine data from the group of subjects. By Gradient Boosting, a model returned an accuracy of 0.81 for classifying Condition B vs. Condition D and 0.77 for Condition B vs. Condition C. Accuracies were over 0.7 for all pairs of conditions. Table 1 shows the results for a four-way classification of conditions, calculated from Experiment 2.

	F1 A	F1 B	F1 C	F1 D	Overall Accuracy
Logistic Regression	0.34	0.25	0.25	0.34	0.20
Support Vector Machine	0.26	0.34	0.11	0.29	0.28
Decision Tree	0.59	0.53	0.62	0.33	0.53
Random Forest	0.68	0.66	0.65	0.63	0.65
Gradient Boosting	**0.81**	**0.71**	**0.72**	**0.71**	**0.75**

Table 1. Machine Learning results from User Acceptance Study data from 17 subjects

Robot Programming

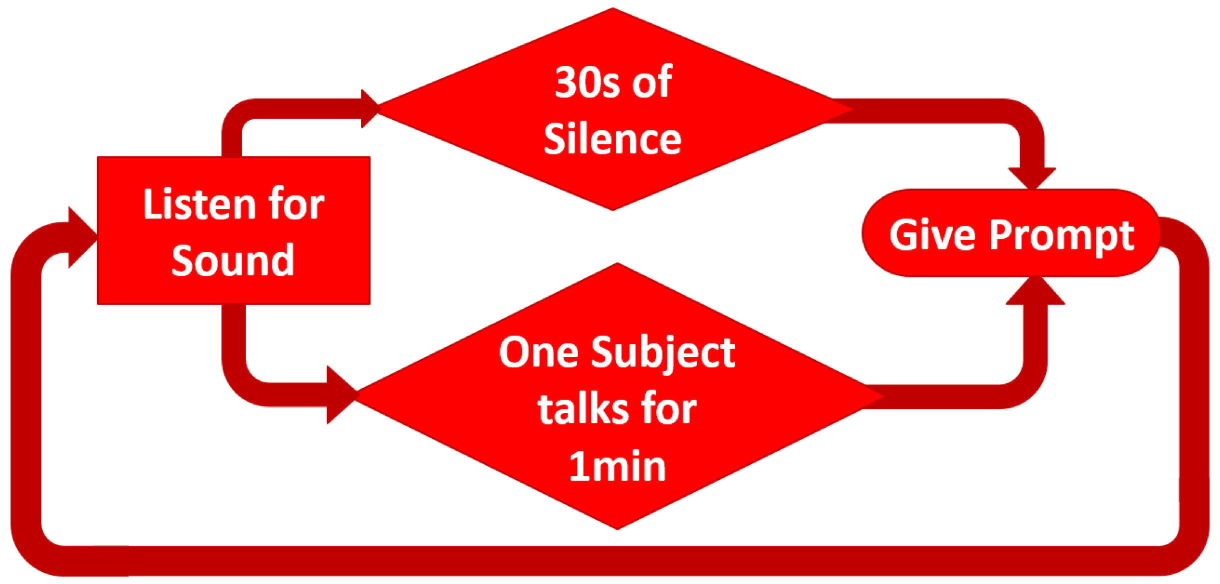

Figure 5. Algorithm that the robot will follow for the Fall 2021 study at the Autism Center

Conclusions

We can draw the following conclusions from the data:

1. The attentiveness of 11 year olds with autism interacting with a robot is correlated with specific physiological responses.
2. Differences in voice type and smoothness of motions during human-robot interactions provoke perceptibly different physiological responses among adults around college age.
3. NAO robots can be programmed to serve as a discussion mediator, detecting pauses or balancing contributions.

References

[1] Tahir, Y., Dauwels, J., Thalmann, D., & Thalmann, N. M. (2018, April 25). A User Study of a Humanoid Robot as a Social Mediator for Two-Person Conversations. International Journal of Social Robotics, 12, 1031-1044.

[2] Welch, K. C., Lahiri, U., Sarkar, N., Warren, Z., Stone, W., Liu, C., & Yildirim, G. (2010). Affect-Sensitive Computing and Autism. In D. Gökçay (Ed.), Affective Computing and Interaction: Psychological, Cognitive and Neuroscientific Perspectives (pp. 325-343). IGI Global

[3] Zhai, J. & Barreto, A. (2006). Stress Recognition Using Non-invasive Technology. In G. Sutcliffe & R. Goebel (eds.), FLAIRS Conference (395-401), : AAAI Press

[4] Sommer, K., Slaughter, V., Wiles, J., Owen, K., Chiba, A. A., Forster, D., Malmir, M., & Nielsen, M. (2021, March). Can a robot teach me that? Children's ability to imitate robots. Journal of Experimental Child Psychology, 203, 105040. https://doi.org/10.1016/j.jecp.2020.105040

Acknowledgements

The authors want to thank the University of Louisville Autism Center, Dr. Ha Do, Rebecca Castelly, Janet Pulgares Soriano, Prasad Nethala, and Shaye Allen for their contributions toward the research.

This research was supported by the National Science Foundation (NSF) under Smart and Connected Health (SCH) Grant #1838808 and REU Grant #1950137.

NSF