Learning to Maintain Engagement: No One Leaves a Sad DragonBot

Goren Gordon and **Cynthia Breazeal**
Personal Robots Group, MIT Media Lab,
20 Ames Street E15-468
Cambridge, MA 02139

Abstract

Engagement is a key factor in every social interaction, be it between humans or humans and robots. Many studies were aimed at designing robot behavior in order to sustain human engagement. Infants and children, however, learn how to engage their caregivers to receive more attention. We used a social robot platform, DragonBot, that learned which of its social behaviors retained human engagement. This was achieved by implementing a reinforcement learning algorithm, wherein the reward is the proximity and number of people near the robot. The experiment was run in the World Science Festival in New York, where hundreds of people interacted with the robot. After more than two continuous hours of interaction, the robot learned by itself that making a sad face was the most rewarding expression. Further analysis showed that after a sad face, people's engagement rose for thirty seconds. In other words, the robot learned by itself in two hours that almost no-one leaves a sad DragonBot.

Introduction

Long-term social interaction is based on continued engagement between participants, either humans or robots. In the human-robot interaction field, there have been several studies that examined and raised the question: what aspects of human and/or robot behavior can sustain engagement (Breazeal and Scassellati 1999; Sidner et al. 2005; Rich et al. 2010)? On the other hand, infants and children learn by themselves how to get and maintain their caregivers' attention (Cohn and Tronick 1987). In this preliminary study we wanted to test whether a social robot can learn by itself which behaviors can sustain people's engagement.

In order to answer this question, we combined principles of developmental robotics (Oudeyer, Kaplan, and Hafner 2007; Gordon and Ahissar 2012) and human-robot interaction. We used the social robotic platform, the DragonBot (Setapen 2012), and implemented a reinforcement learning algorithm (Degris, White, and Sutton 2012) such that the robot seeks which behavior maximizes its future accumulated rewards. Here, the rewards were proportional to the engagement of people near the robot, measured by the proximity and number of people facing the robot.

Experimental Setup

DragonBot is a very expressive social robotic platform and has a large repertoire of possible facial expressions and actions (Setapen 2012). We chose nine non-verbal behaviors that served as the action-space of the sought-for policy: Yes (small nod, 'mm' utterance); I like it (nod, 'alright' utterance); Mmmmm (looking up, 'Mmmm' utterance); mmHmm (looking up, 'mmHmmm' utterance); Sad (moving down, move eyebrows, sad sound); Shy (moving down, big eyes, embarrassed laugh); Think (lower lid, 'mmm' utterance); Yawn (large movement upward, yawning sound); Mph (moving down, lower eyelids, 'Mph' utterance).

For sensing, a camera above the robot supplied a visual image, whereupon we used a standard face-recognition algorithm to find faces in that image, i.e. it only recognized people facing the robot. The state-space was thus composed of a discretized measure of the distance of the most proximal recognized face (near, medium ,far) and the previous action, to allow for complex behaviors.

The reward was set to be proportional to the sum of the faces' sizes, i.e. the closer people were and more numerous, the larger the reward. This reward function was selected to approximate people's engagement, on the basis that people facing the robot in close proximity are in the act of engagement with the robot.

We run the experiment in the World Science Festival (WSF) in New York on June 1st 2014. This challenging environment embedded large "competition" for attention for our DragonBot, as there were many other robots in the same hall. Nevertheless, this opportunity resulted in hundreds of people interacting with the robot in the course of several hours.

Results

Due to the close quarters of the WSF environment, and the relatively short duration of data collection (2.3 hours), analysis of the resulting policy is restricted to summation over the state-space, i.e. we report here only on the specific behaviors the DragonBot learned, and not more complex and state-specific policies. Figure 1 shows the action-values learned over time. One action becomes more valuable than the rest: the Sad behavior. This shows that the robot learned, by itself, that making a sad face results in more people staying for

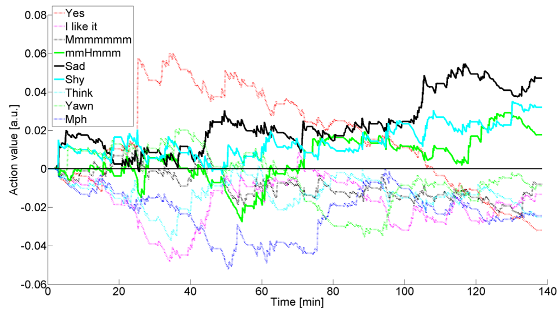

Figure 1: Policy-learning dynamics, measured as the action-value of each action as a function of time. Action-value is calculated as $V(a) = \sum_s Q(s, a)$. Dotted lines denote actions that ended with action-value below zero, for visualization purposes.

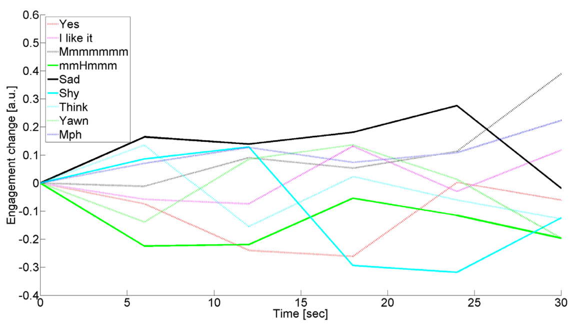

Figure 2: Engagement change as a function of time after each behavior's execution, measured as the difference in reward from the onset of the behavior.

longer, i.e. maintained engagement. Notice that the second most valuable behavior is the Shy behavior, also employing a large-eye paedomorphic expression.

We next analyzed the effects of each behavior on engagement, measured by the change in reward, i.e. we time-aligned all the behaviors and measured the change in the number and proximity of detected faces following each behavior. Figure 2 shows that during the first 30 seconds after the behavior is performed, the Sad behavior dominates with positive engagement change, i.e. on average people did not leave the robot for half a minute after a sad face. While these results are preliminary and not significant, they indicate a possible trend in people's attitude, which the robot learned via our algorithm.

These results supports the personal impression of the experimenter in the WSF (GG), where regardless of age, gender and race, whenever the DragonBot made a sad face, everyone mimicked it with empathy, thereby staying near it, whereas the other behaviors had mixed responses.

Conclusions and Future Work

This preliminary study, although conducted in a real-world and challenging environment, showed that a social robot can learn by itself, relatively quickly, which non-verbal behaviors maintain people's engagement, at least for a short while. An infant-like face, with "big sad eyes" has been shown to influence human behavior (Waller et al. 2013). We suggest that DragonBot may have learned this psychological aspect of human emotional response via a developmental-robotics inspired approach. It learned that "no leaves a sad DragonBot".

More research is required to further validate and extend these preliminary results to more complex scenarios, such as responses to humans' facial expressions, including more complex sequences of behaviors and including verbal communication. Nevertheless, we believe that the first steps towards a socially learning robot, one that learns by itself how to interact with people, are finally on their way.

Acknowledgments

G.G. was supported by the Fulbright commission for Israel, the United States-Israel Educational Foundation. This research was supported by the National Science Foundation (NSF) under Grants CCF-1138986. Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not represent the views of the NSF.

References

Breazeal, C., and Scassellati, B. 1999. How to build robots that make friends and influence people. In *Intelligent Robots and Systems, 1999. IROS '99. Proceedings. 1999 IEEE/RSJ International Conference on*, volume 2, 858–863 vol.2.

Cohn, J. F., and Tronick, E. Z. 1987. Mother-infant face-to-face interaction: The sequence of dyadic states at 3, 6, and 9 months. *Developmental Psychology* 23(1):68–77.

Degris, T.; White, M.; and Sutton, R. S. 2012. Off-policy actor critic. In *ICML2012*.

Gordon, G., and Ahissar, E. 2012. *A Curious Emergence of Reaching*, volume 7429 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. book section 1, 1–12.

Oudeyer, P. Y.; Kaplan, F.; and Hafner, V. V. 2007. Intrinsic motivation systems for autonomous mental development. *Evolutionary Computation, IEEE Transactions on* 11(2):265–286.

Rich, C.; Ponsler, B.; Holroyd, A.; and Sidner, C. L. 2010. Recognizing engagement in human-robot interaction. In *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, 375–382.

Setapen, A. 2012. *Creating Robotic Characters for Long-term Interaction.* Thesis.

Sidner, C. L.; Lee, C.; Kidd, C. D.; Lesh, N.; and Rich, C. 2005. Explorations in engagement for humans and robots. *Artificial Intelligence* 166(1):140–164.

Waller, B.; Peirce, K.; Caeiro, C.; Scheider, L.; Burrows, A.; McCune, S.; and Kaminski, J. 2013. Paedomorphic facial expressions give dogs a selective advantage. *PLoS ONE* 8(12):e82686.