The Effect of Duration for Human Feedback to Robot in Reinforcement Learning

Thuy Vy Tran
Tufts University Department of Computer Science
The Assistive Agent Behavior and Learning Lab (AABL)

***Abstract*—Most studies focus on the task at hand for the robot and how human feedback affects it in reinforcement learning. In reinforcement learning focuses on how the agent (robot) might react in an environment (human feedback) when given a reward. It is important to look at reinforcement learning with a duration of time in mind. We will look at how time affects human feedback when teaching the robot could get a different result. Also, it depends on the difficulty of the task and time.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

In interactive reinforcement learning (RL), the robot (the agent) gets human feedback (the environment) while learning the task at hand. It is a crucial tool that allows researchers to look at the optimal behavior of the robot through trial-and-error with the environment.[4] However, the human feedback is not perfect at all times. The imperfection of the teaching affects how robots learn the task.

The causes of imperfect human feedback could range from boredom to fatigue as an effect changes the outcome of the reward given to the robot. [2] A real-life example of human error would be when a chef adds too much of an ingredient to their dish. The outcome of the cooking will change due to this imperfection. The comparison is similar to how a mistake from the human teacher can make the robot learn at a slower rate or not complete the task correctly.In our study, we will be looking at the effect fatigue has on the duration of time in HRI.

II. Background

There have been studies that show the effect fatigue has on work. The more fatigue the human is, the quality of work produced goes down. [5] The longer the duration the human gives feedback to a robot, the more likely fatigue will set in. In return affects how the outcome of the robot learning in RL.

However, robots with a longer duration of time also adjust the learning curve of a task. I have not found many studies that focus on the relationship of time in RL for HRI. At some point, there is going to be a time where there will be no human feedback. An example would be when humans are taking a break. The robot can learn from their environment during that time.[1] However, inaccurate feedback can still occur during this time.

III. Methodology

The focus of the research is to analyze if changes in human behavior affect robot performance over time. I will be using TAMER framework to analyze the study.[3] The TAMER framework uses the feedback (human) given to the agents(the robots) as a reward system. One of the benefits is no technical knowledge is needed to analyze the relationship. TAMER allows me to analyze the variable behind why human feedback over time without worrying if the user is familiar with robots or not.

IV. User Study

The study will be conducted online by using OpenGym AI two programs, Taxi and FrozenLakes. A pre-survey is given before they start the task. After an increment of 10, 30 minutes, and an hour a questionnaire will be handed to the user. Also, the percentage of how close they are to finishing the goal is given during this time. In the end, they will be handed the same pre-survey to complete.

V. Future Works

I would have loved to get more into the programming and algorithmic part of my research. Another thing I would love to have completed is the result, discussion, and conclusion. I learned how to read and analyze research articles for my research. I also learned the writing style and rules for research. Being able to learn the libraries and interfaces used in robotics helped me understand more about machine learning.

References

[1] Taylor Kessler Faulkner, Reymundo A Gutierrez, Elaine Schaertl Short, Guy Hoffman, and Andrea L Thomaz. [n. d.] Active Attention-Modified Policy Shaping. en, 9.

[2] W. Bradley Knox and Peter Stone. 2010. Combining manual feedback with subsequent mdp reward signals for reinforcement learning. In *In AAMAS*.

[3] W. Bradley Knox, Peter Stone, and Cynthia Breazeal. 2013. Teaching agents with human feedback: a demonstration of the TAMER framework. In *Proceedings of the companion publication of the 2013 international conference on Intelligent user interfaces companion* (IUI '13 Companion). Association for Computing Machinery, New York, NY, USA, (March 2013), 65–66. ISBN: 978-1-4503-1966-9. DOI: 10.1145/2451176.2451201. Retrieved 06/17/2021 from https://doi.org/10.1145/2451176.2451201.

[4] Jens Kober, J. Andrew Bagnell, and Jan Peters. 2013. Reinforcement learning in robotics: A survey. en. *The International Journal of Robotics Research*, 32, 11, (September 2013), 1238–1274. Publisher: SAGE Publications Ltd STM. ISSN: 0278-3649. DOI: 10.1177/0278364913495721. Retrieved 08/03/2021 from https://doi.org/10.1177/0278364913495721.

[5] Sara Viotti, Lynnette Essenmacher, Lydia E. Hamblin, and Judith E. Arnetz. 2018. Testing the reciprocal associations among co-worker incivility, organisational inefficiency, and work-related exhaustion: A one-year, cross-lagged study. *Work & Stress*, 32, 4, (October 2018), 334–356. ISSN: 02678373. DOI: 10.1080/02678373.2018.1436615. Retrieved 06/24/2021 from http://mutex.gmu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=pbh&AN=132293333&site=ehost-live.