Human-agent, Co-operative Reinforcement Learning

UMass Team:

Yash Chandak (Ph.D. Mentor)
Aarshee Mishra
Abhik Singla
Nicholas Perello
William Goss

Microsoft Team:

Soundar Srinivasan
Aalesha Rajaram
Vivek Chettia
Vishal Rohra
H M Sajjad Hossain

Outline

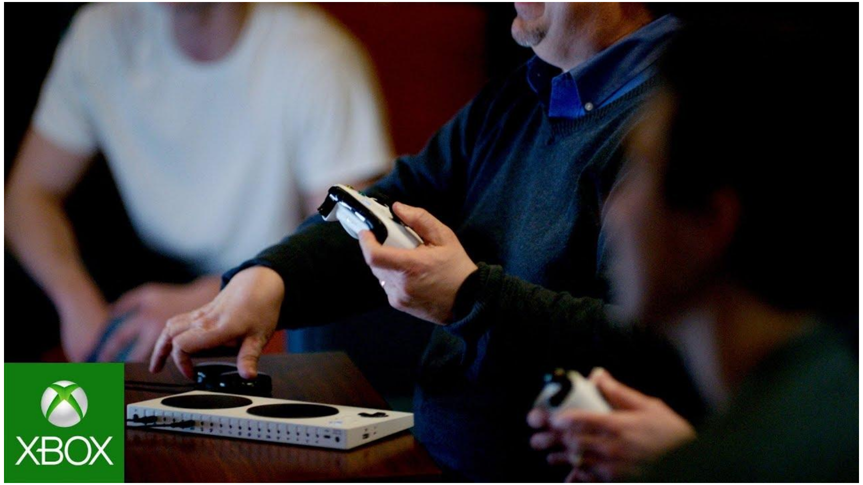
- Introduction
- Challenges
- Related Works
- Proposed Approach
- Experiments
- Contingency Plan

Introduction

How can we make experiences, such as gaming, more inclusive to everyone of all abilities?

Microsoft has introduced the Xbox Adaptive Controller to help meet the needs of gamers with limited mobility.

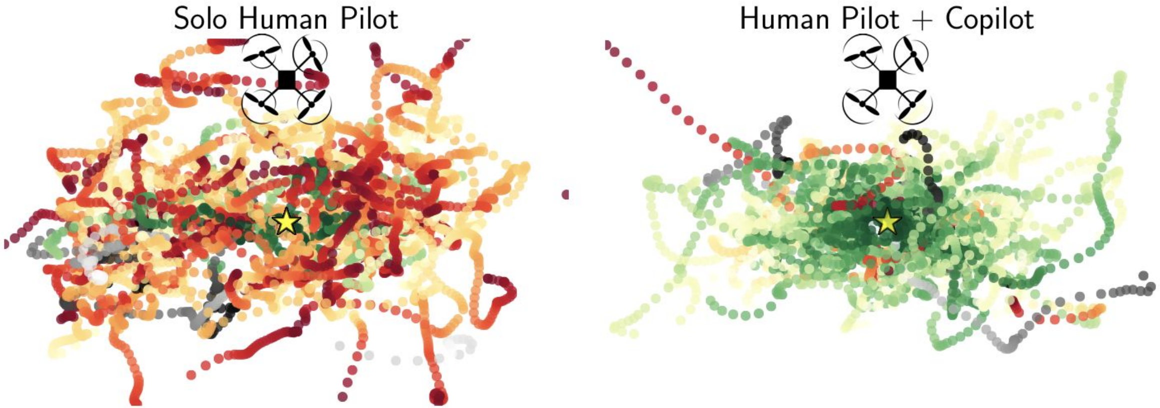
Can this be further improved using reinforcement learning?



Introduction

We aim to investigate human-in-the-loop reinforcement learning in a multi-agent co-operative environment.

In human-in-the-loop RL, the virtual agent can assist the human to complete a common task.



A bird's-eye view of trajectories followed by human pilots with and without a copilot on the quadrotor landing task.

Challenges

Needs to be adaptive to the human player's intent so that it does not destroy the player’s experience.

How to define and identifying a user’s intent and their ability.

Intervention - When and to what extent?

Gathering human player data.

Related work

On the Utility of Learning about Humans for Human-AI Coordination: In training for cooperative games using AI, we should include human models.

Shared Autonomy via Deep RL: Shared Action space based end-to-end model free human in the loop RL framework

Maximum Entropy Inverse Reinforcement Learning: Estimating reward function by observing human trajectories.

Continuous Control with Deep RL: A deep RL framework for solving continuous control problems

Cooperative Inverse RL: Model based framework where the agent knows that it is in a shared environment and attempts to maximize the human's reward.

Human-level performance in 3D multiplayer games with population-based RL: Training a huge number of agents concurrently to achieve human level performance in FPS games.

An Efficient, Generalized Bellman Update For Cooperative IRL: Exponential reduction in the size of action space for CIRL.

Proposed Approaches

Shared Action Space:

Agent pre-trained on the environment

Execute human policy if action value is within 1-α of the optimal action value

Disjoint Action Space:

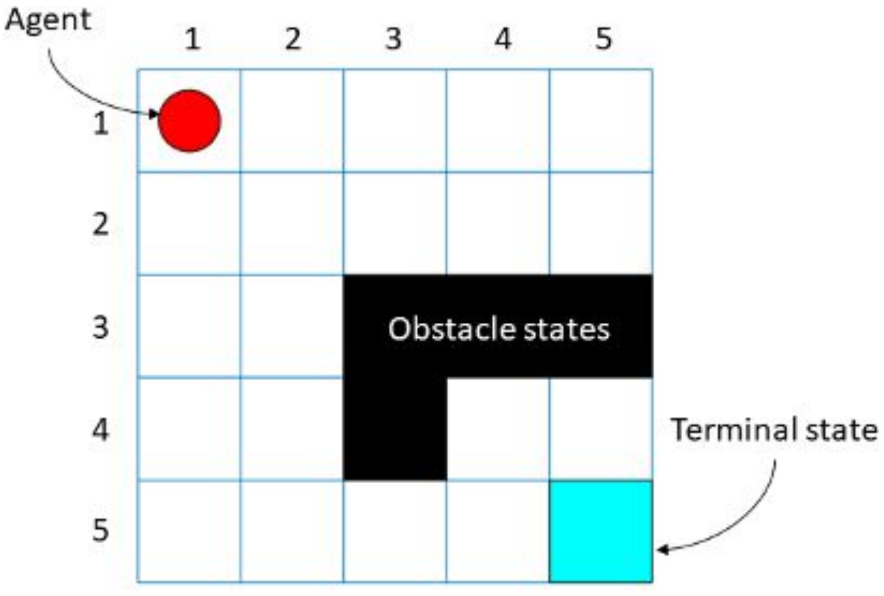
Behaviour cloning human model as part of action space

Train the agent on this new environment

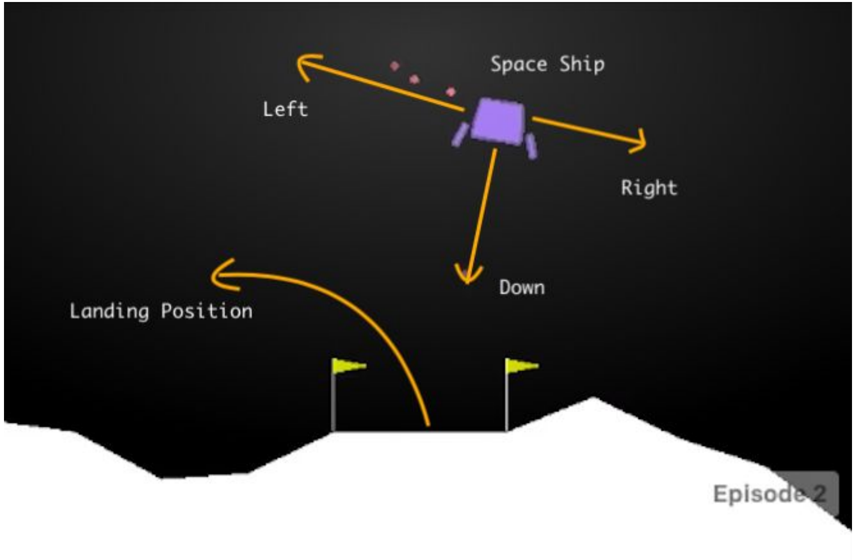
Figuring out the intent of the human as a new reward function R_{hidden} apart from the usual reward using inverse RL

Experiments

Baseline: Achieve the highest score in collaboration, ignoring the fun part.



Grid World



Lunar Lander

Timeline until Midpoint

2/14 - Finalizing literature review and begin design

2/21 - Design and initial implementation

2/28 - Implementation and experimenting on a simple environment

3/6 - Experimenting on an altered simple environment

3/13 - Experimenting on an advanced environment

3/20 - *No Class - Spring Vacation*

3/27 - **Midpoint** presentations (Two working prototypes)

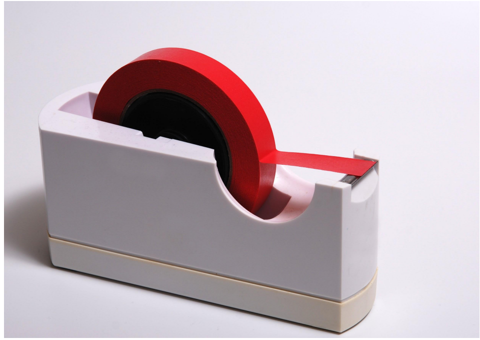
Challenges for Completion

Humans can have varying levels of tolerance

Human model could be flawed

Evolution of training

User studies = Time + Paperwork



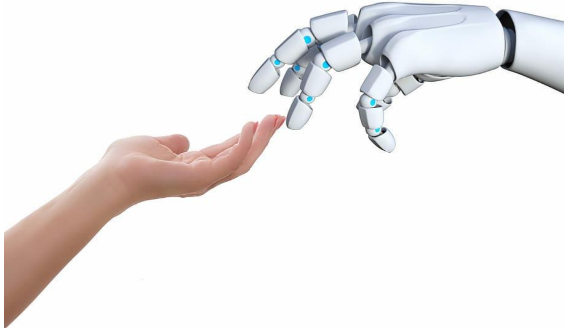
Contingency Plan

Ignoring user intent

Complete take over

Useful additions from information we gain along the way

- When to have the agent intervene?
- How long should it intervene?
- To what degree should the agent act?



Questions?