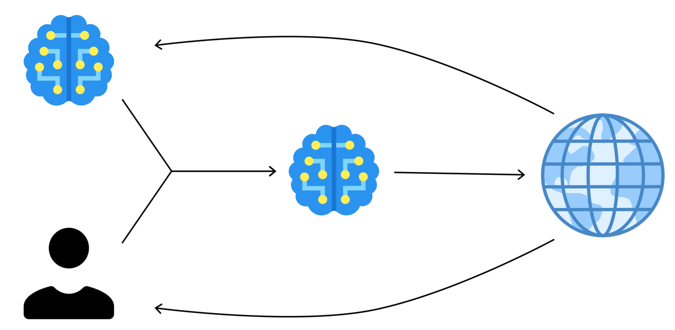
Demonstrating Optimized Delegation between AI and Human Agents

Andrew Fuchs[1,2], Andrea Passarella[1,2], Marco Conti[2]

[1]Università di Pisa, Pisa, Italy.
[2]Istituto di Informatica e Telematica (IIT) - CNR, Pisa, Italy.

Introduction

With humans interacting with AI-based systems at an increasing rate, it is necessary to ensure the artificial systems are acting in a manner which reflects understanding of the human. We note the significance of comprehension and response to the actions or capabilities of a human from an agent's perspective, as well as the possibility to delegate decisions either to humans or to agents, depending on who is deemed more suitable for a given context. To that end, we investigate the use of cognitively inspired models of behavior to predict the behavior of both human and AI agents. The predicted behavior and performance is used to delegate control between humans and AI agents through the use of an intermediary manager entity. As we demonstrate, this allows overcoming potential shortcomings of either humans or agents in the pursuit of a goal.



Key Concepts

Our scenario is comprised of the following key concepts:

- Team of potentially erroneous actors
  - Example: Human working with AI/RL agent
- Cognitively-inspired model of human behavior
- Manager delegating control
- Optimized delegation of actions

Key Scenario

The key aspects of the scenario and task we are trying to achieve:

- Mixture of Q-Learning [1] and Instance-Based Learning (IBL) [2, 3] agents trained to navigate environment
- Artificial errors imposed on navigating agents
  - Errors are sub-optimal actions w.r.t. shortest path
  - Each agent has predetermined probability of error in error states
- Use of different models guiding behavior to demonstrate variance in team members

Desired Outcomes

Train a manager which:

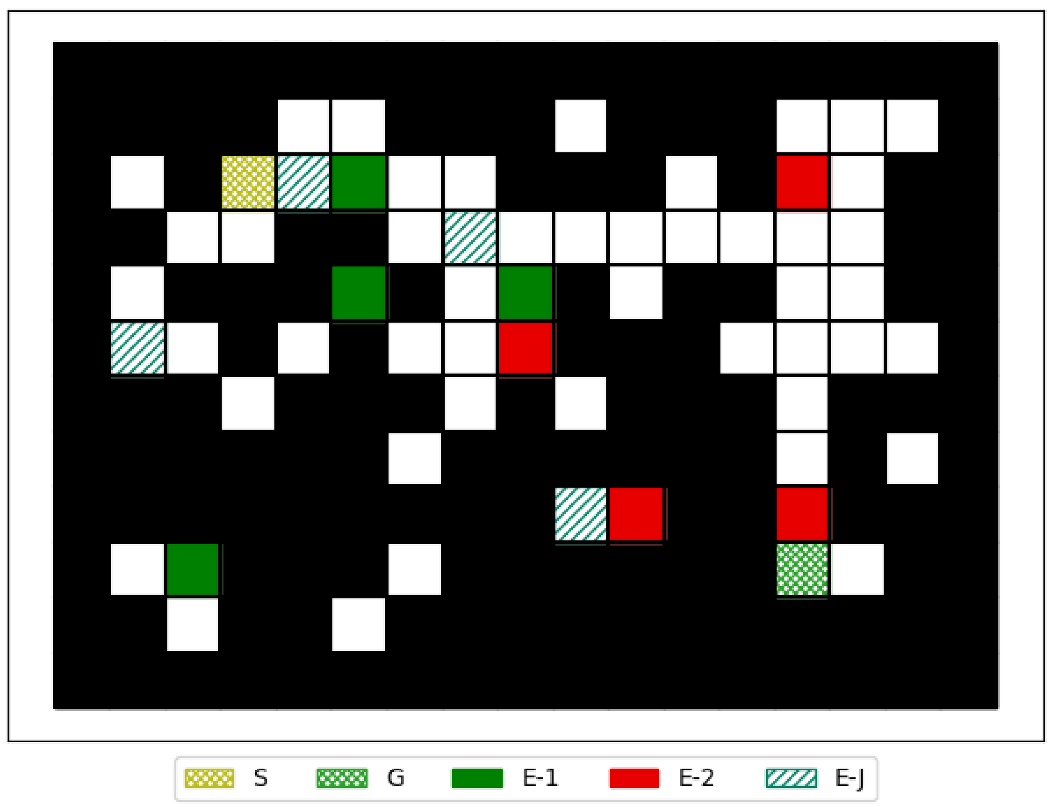
- Learns the ideal candidate for action via observations of behavior
- Demonstrates ability to predict/anticipate performance of human-like behavior
- Enables teams outperforming individual performance or random selection of team members

Environment and Agents

To test and demonstrate performance for the task, we utilized the gridworld scenario.

Environment

A modified gridworld with error states inserted:



with "E-1" and "E-2" referring to error states for agent_1 and agent_2 respectively, and "E-J" denoting an error state for all agents.

- Parameters:
  - Wall cell %
  - Start/goal position
  - Frequencies of error states

Agents

- Grid navigating agents:
  - We use a mixture of agent types
  - Q-Learning represents an AI agent

$$Q_{t+1}(s,a) = (1-\alpha)Q_t(s,a) + \alpha[r + \gamma \max_{a'} \{Q_t(s',a')\}]$$

  - IBL represents human-like learning/reasoning

$$V(s,a) = \sum_{i=1}^{n} p_i x_i$$

  - Navigating agents select movement action in gridworld
  - Rewards are based on movement actions and game outcomes
- Manager agent:
  - Uses IBL representation for learning a model of behavior
  - Manager only selects agents, not the movement actions
  - Rewards are based on game outcomes, not movement actions of selected agents

Performance Measures

With our teams and scenarios, we focused on the following aspects of performance:

Agent vs. Team

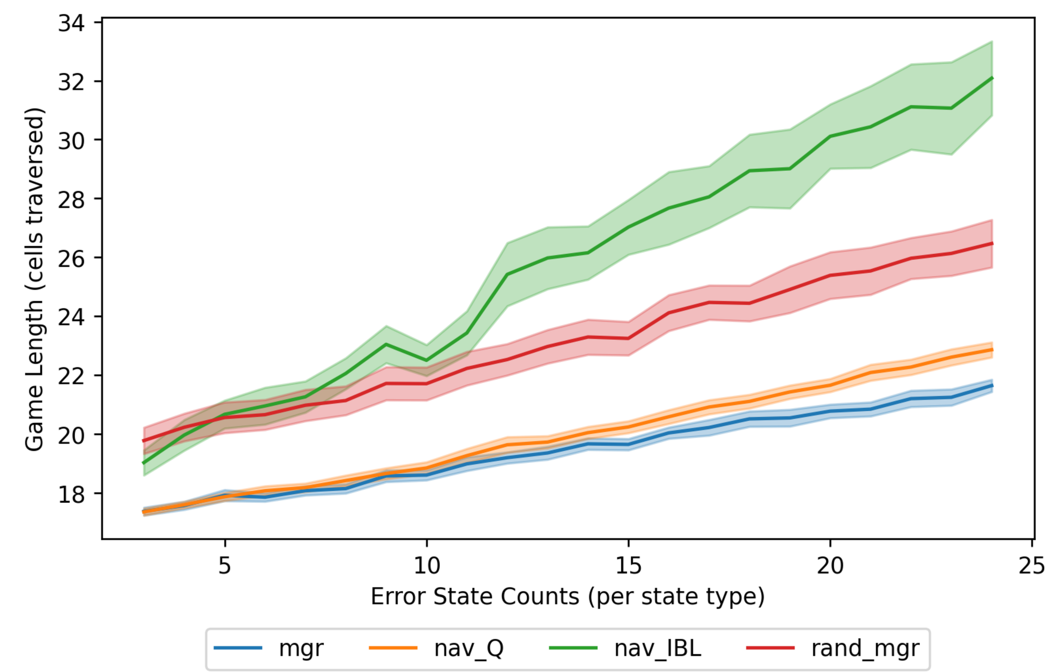
- Compare path lengths for agents operating solo or in a team

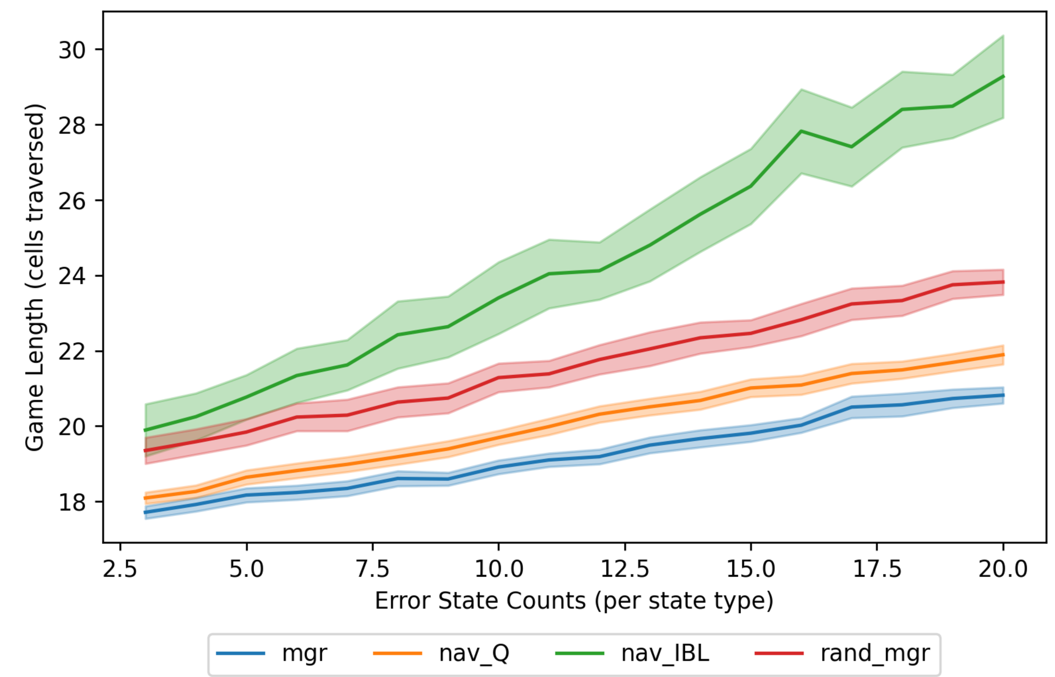
Team Comparisons

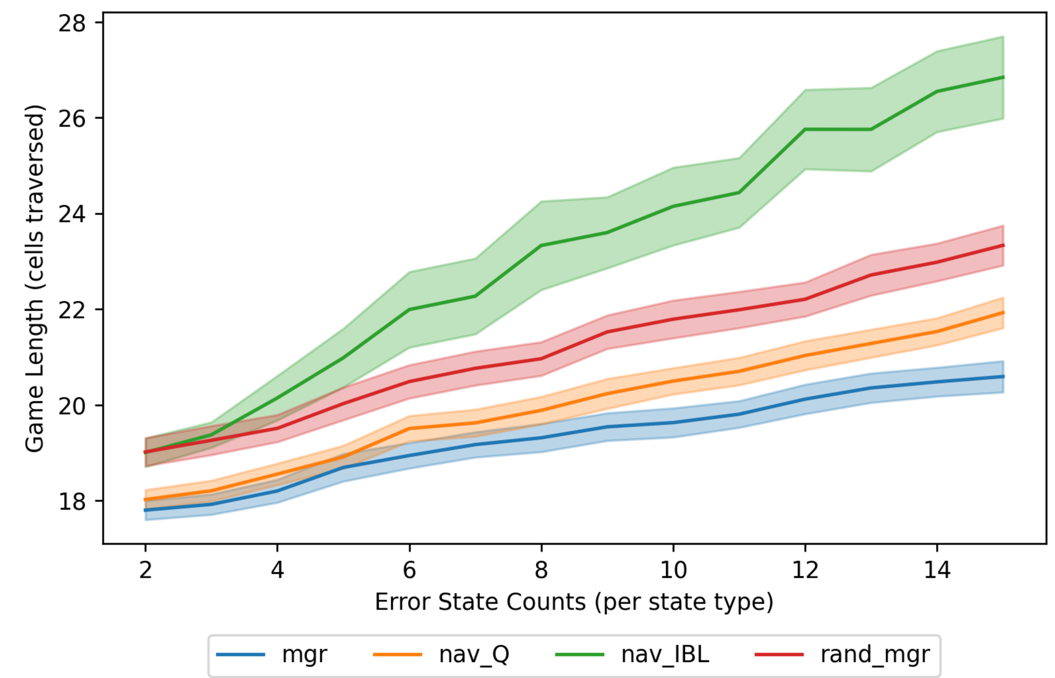
- Compare manager vs random agent selection
- Measure preference for particular agents and how that corresponds to error rates

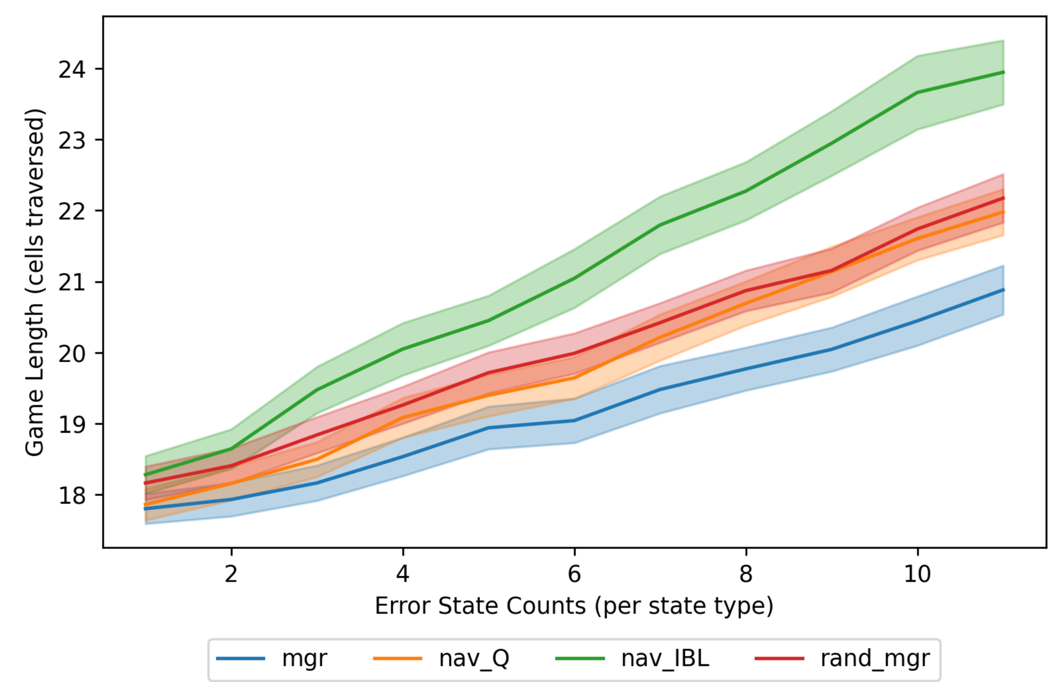
Results

We compare the performance of solo agents, managed agents, and randomly selected agents with respect to game length:









Conclusions

As is demonstrated in our results, the use of a manager agent improves the team performance over random selection (as expected) and can significantly overcome the performance of the agents operating individually. Through observations of agent behavior, the manager learns which agent would best behave given the current state of the environment. Further, manager agents showed a stochastic policy which allowed for agents to be selected with frequency proportional to their likelihood of correct action selection.

References

[1] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction.* MIT press, 2018.

[2] Thuy Ngoc Nguyen and Cleotilde Gonzalez. Cognitive Machine Theory of Mind. 2020.

[3] Thuy Ngoc Nguyen and Cleotilde Gonzalez. Effects of Decision Complexity in Goal-seeking Gridworlds: A Comparison of instance-based learning and reinforcement learning agents. 2020.

andrew.fuchs@phd.unipi.it