Designs for Enabling Collaboration in Human-Machine Teaming via Interactive and Explainable Systems

Rohan Paleja[1], Michael Munje[2], Kimberlee Chestnut Chang[1], Reed Jensen[1], and Matthew Gombolay[3]
[1]MIT Lincoln Laboratory, [2]University of Texas at Austin, [3]Georgia Institute of Technology

Neural Information Processing Systems (NeurIPS)
(December 2024)

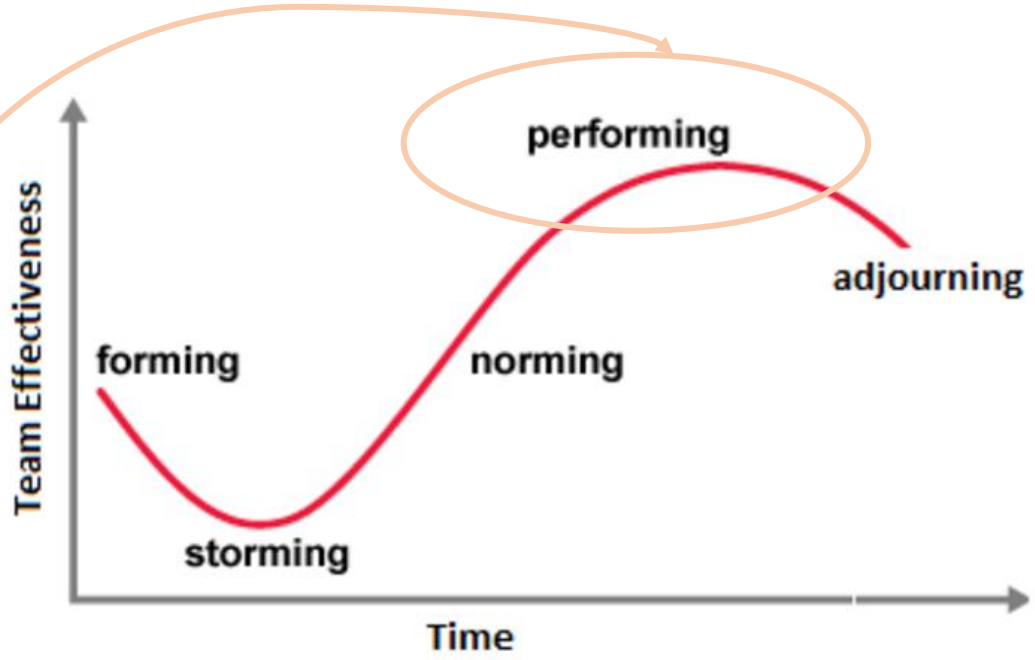
Motivation

A Dynamic and Development Interaction Between Humans and AI

Humans, when teaming with machines, should be able to intuitively update what the robot has learned or change it based upon preferences that evolve over time.

Human team development proceeds through several stages before achieving maximal performance [Tuckman, 1965]

How can we facilitate human-robot teams to reach this stage?

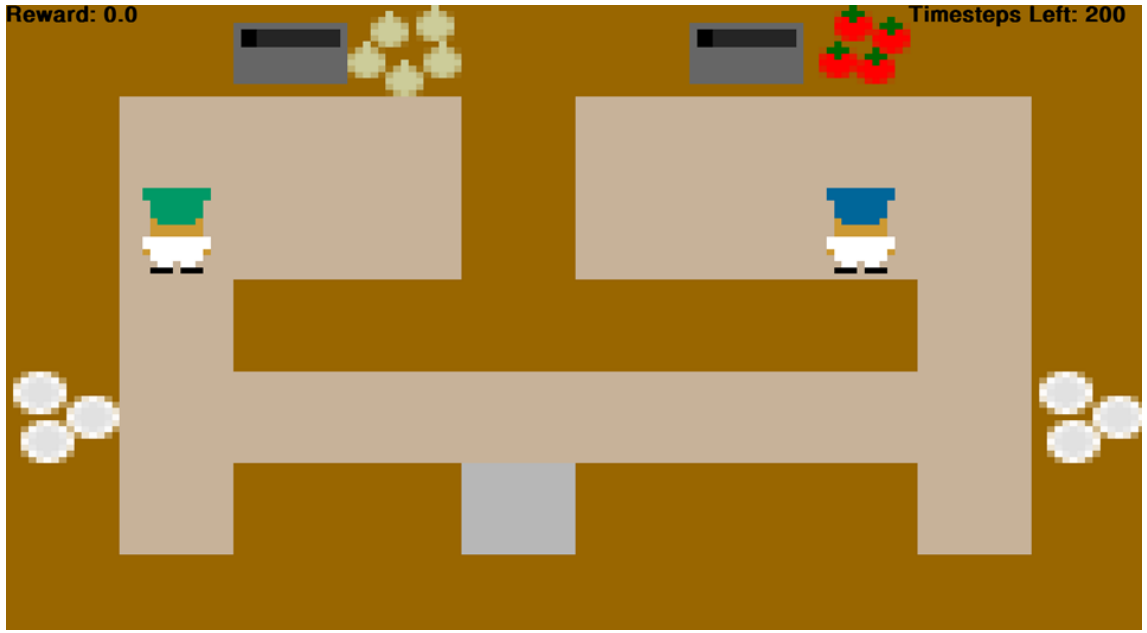


[3] Tuckman, B. W. (1965). Developmental sequence in small groups.. Psychological Bulletin, 63, 384--399.

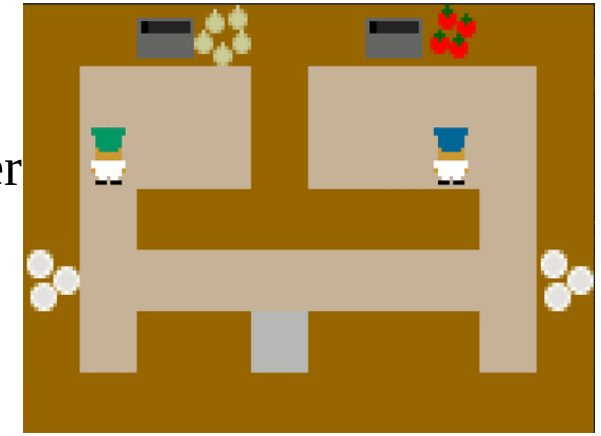
Case Study on Prior HMT Frameworks

Fictitious Co-Play [Strouse et al.]

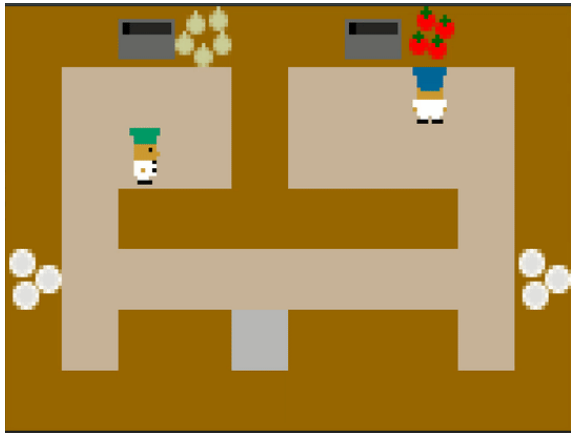
- Fictitious Co-Play (Strouse et al.)
  - Trains with a population of diverse synthetic partners to create an AI that can collaborate with diverse-skilled human players



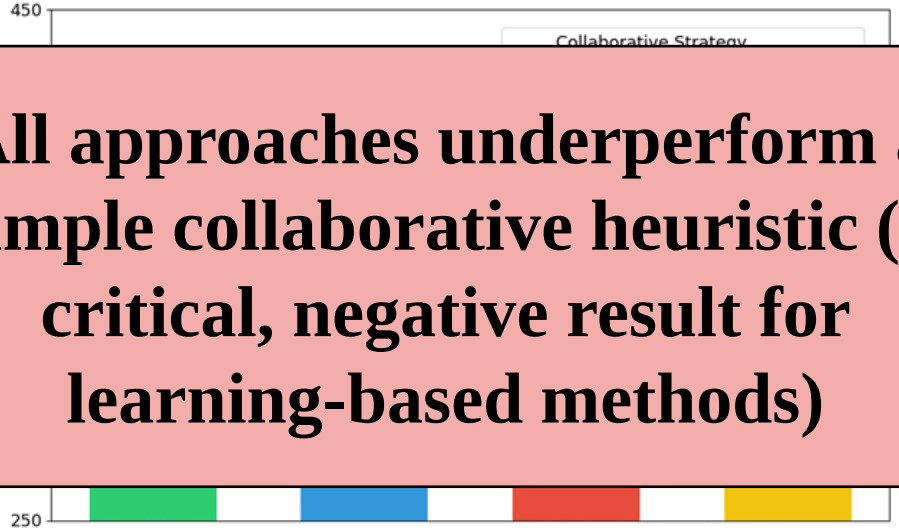
Optional Collaboration



Reward: 306



Reward: 408

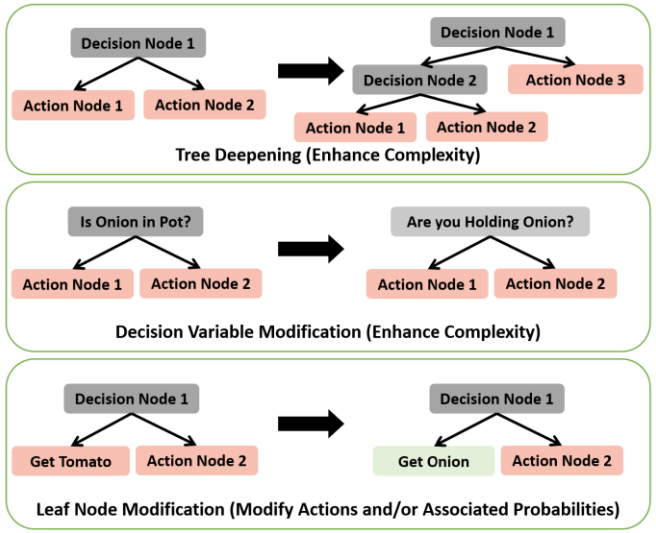


All approaches underperform a simple collaborative heuristic (a critical, negative result for learning-based methods)

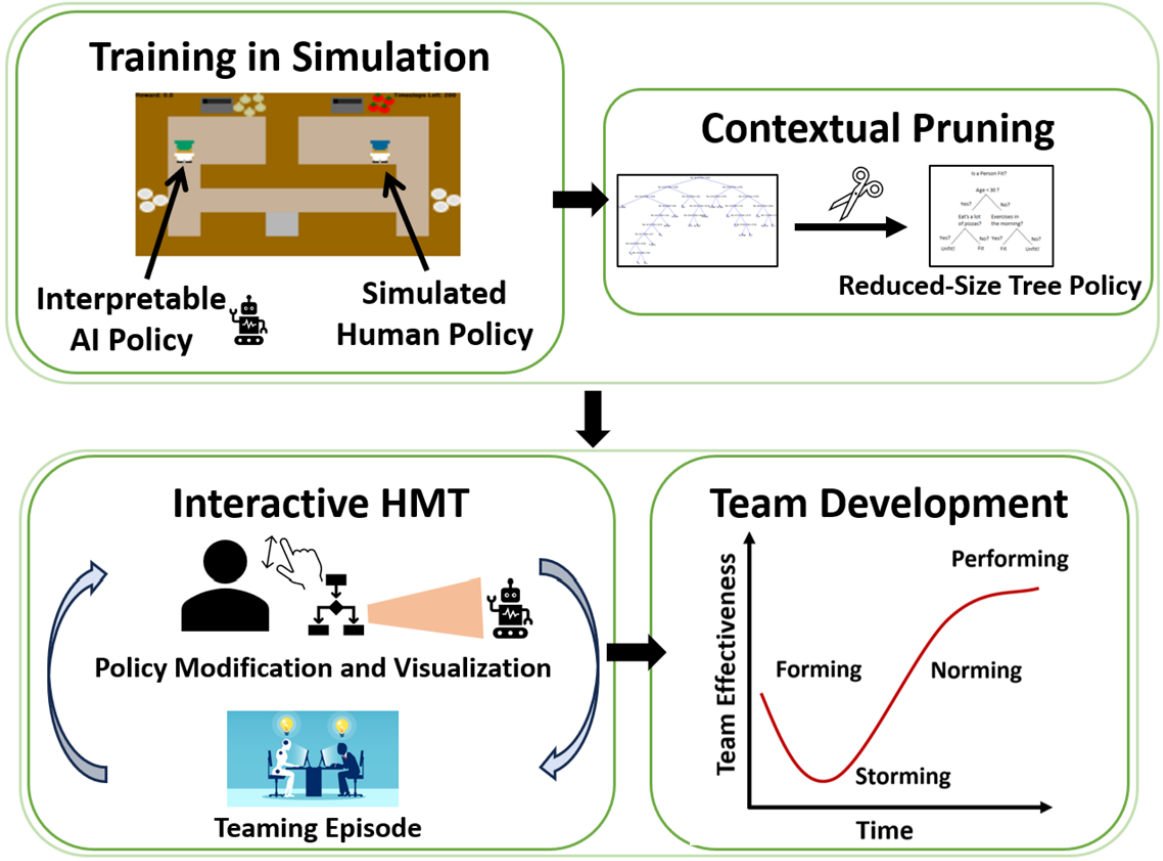
Strouse, D., McKee, K.R., Botvinick, M.M., Hughes, E., & Everett, R. (2021). Collaborating with Humans without Human Data. ArXiv, abs/2110.08176.

Our Proposed Solution: Human-Led Policy Modification

We start by training two separate agents jointly via single-agent PPO.



Users have several capabilities in creating an effective teammate



After training, we deploy our tree-based model in a human-subjects study. Here, users repeatedly team with an AI and interactively reprogram their interpretable AI teammate.

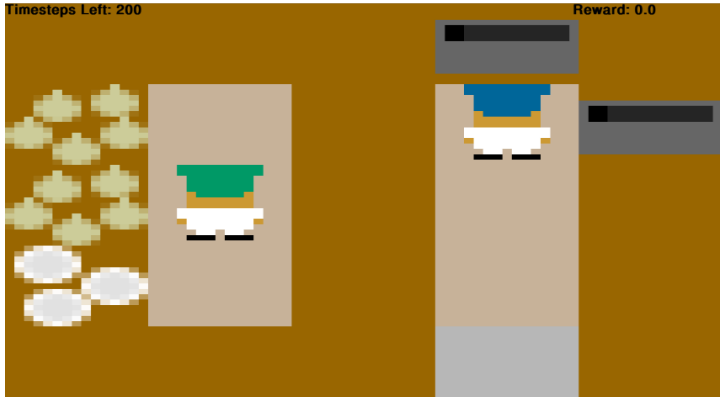
We design a post-hoc ***contextual pruning*** algorithm that allows us to simplify large IDCT models while precisely adhering to model behavior by accounting for:

- Node Hierarchy
- Impossible Subspaces of the State Space

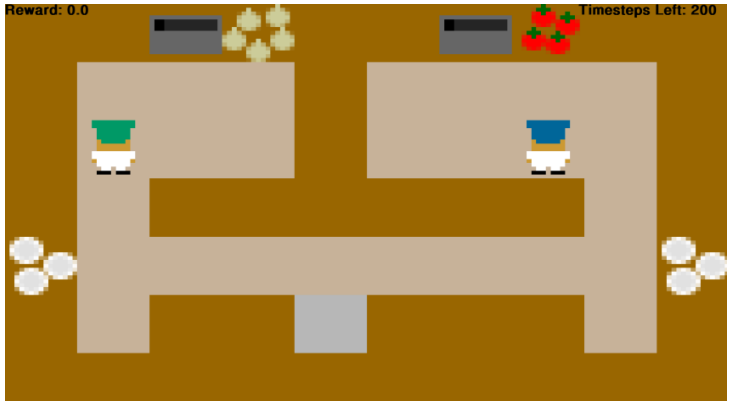
As subjects perform this process repeatedly, we study how team performance changes over time and relate this to Tuckman's Model of Team Development.

Human-Subjects Study

- RQ1: How does team coordination performance vary across different factors?
- RQ2: How does team development vary across different factors?



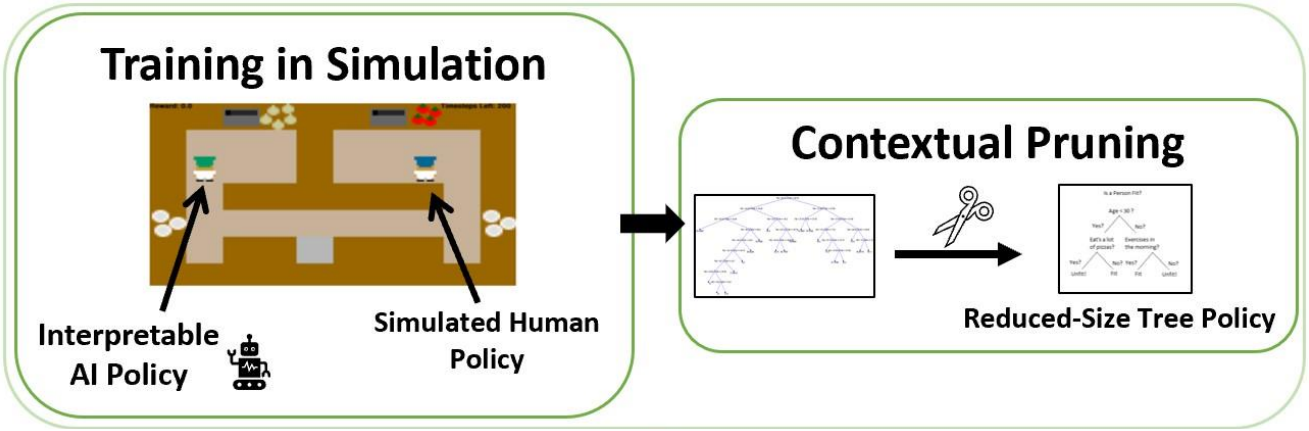
Forced Coordination



Optional Collaboration

Carroll, M., Shah, R., Ho, M.K., Griffiths, T.L., Seshia, S.A., Abbeel, P., & Dragan, A.D. (2019). On the Utility of Learning about Humans for Human-AI Coordination. ArXiv, abs/1910.05789.

Creating Interpretable AI Teammates



InterpretableML Architecture: Interpretable Discrete Control Tree (IDCT)

The resultant representation after training is that of a simple decision tree with categorical probability distributions at each leaf node

State Space

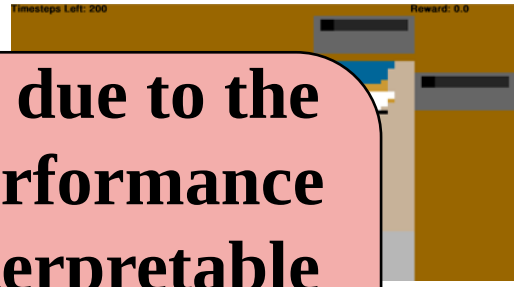
- R

A

-

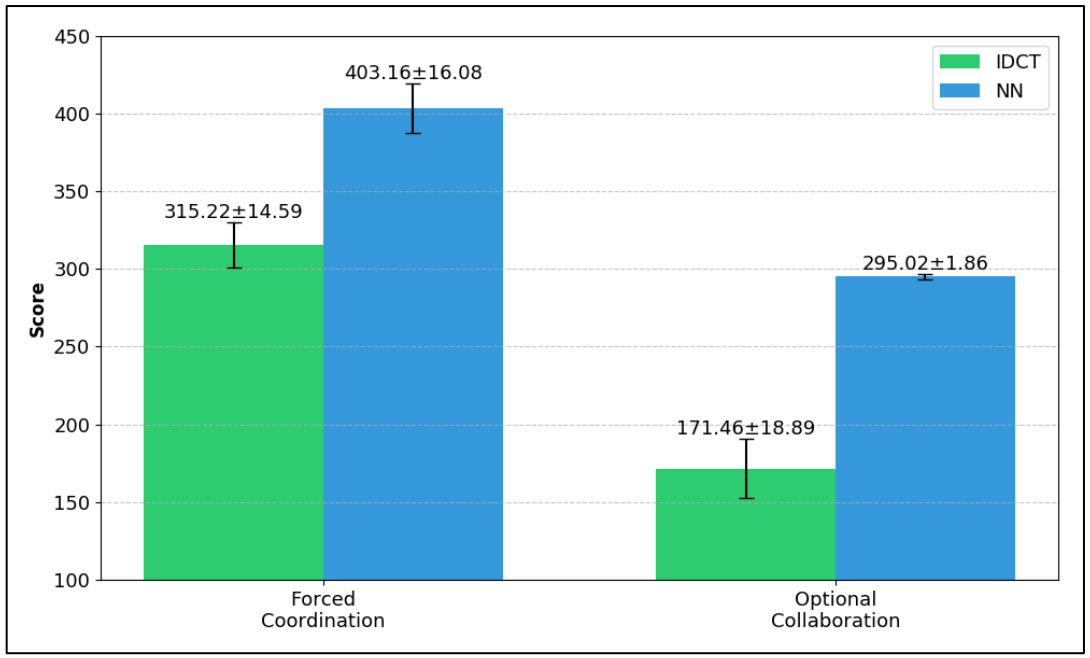
A
re

A consequent confound due to the current difference in performance capabilities between interpretable vs. black-box models is that the NN policy outperforms the IDCT policy in both domains.



nation

Optional Collaboration

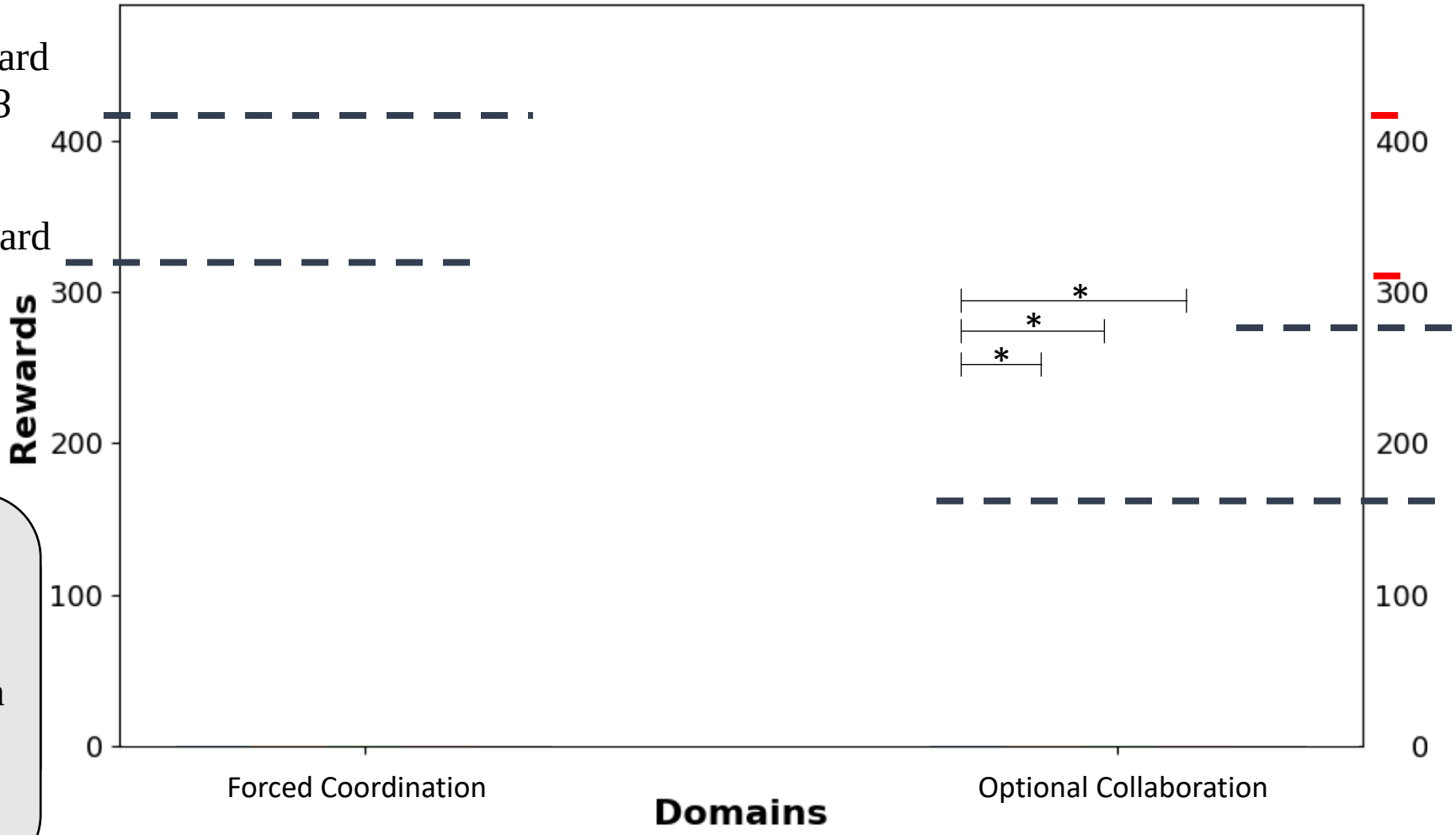


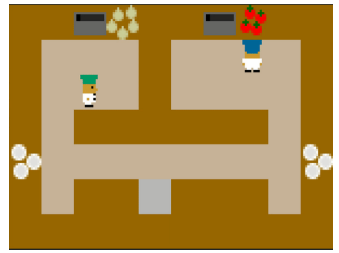
Original Model Size is 256 Leaves

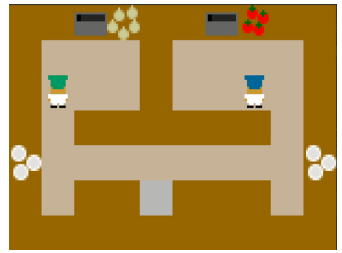
Team Coordination Performance

NN Training Reward
403.16 ± 16.08

IDCT Training Reward
315.22 ± 14.59







NN Training Reward
295.02 ± 1.86

IDCT Training Reward
171.46 ± 18.89

White-box approaches supported with policy modification can outperform white-box approaches alone.

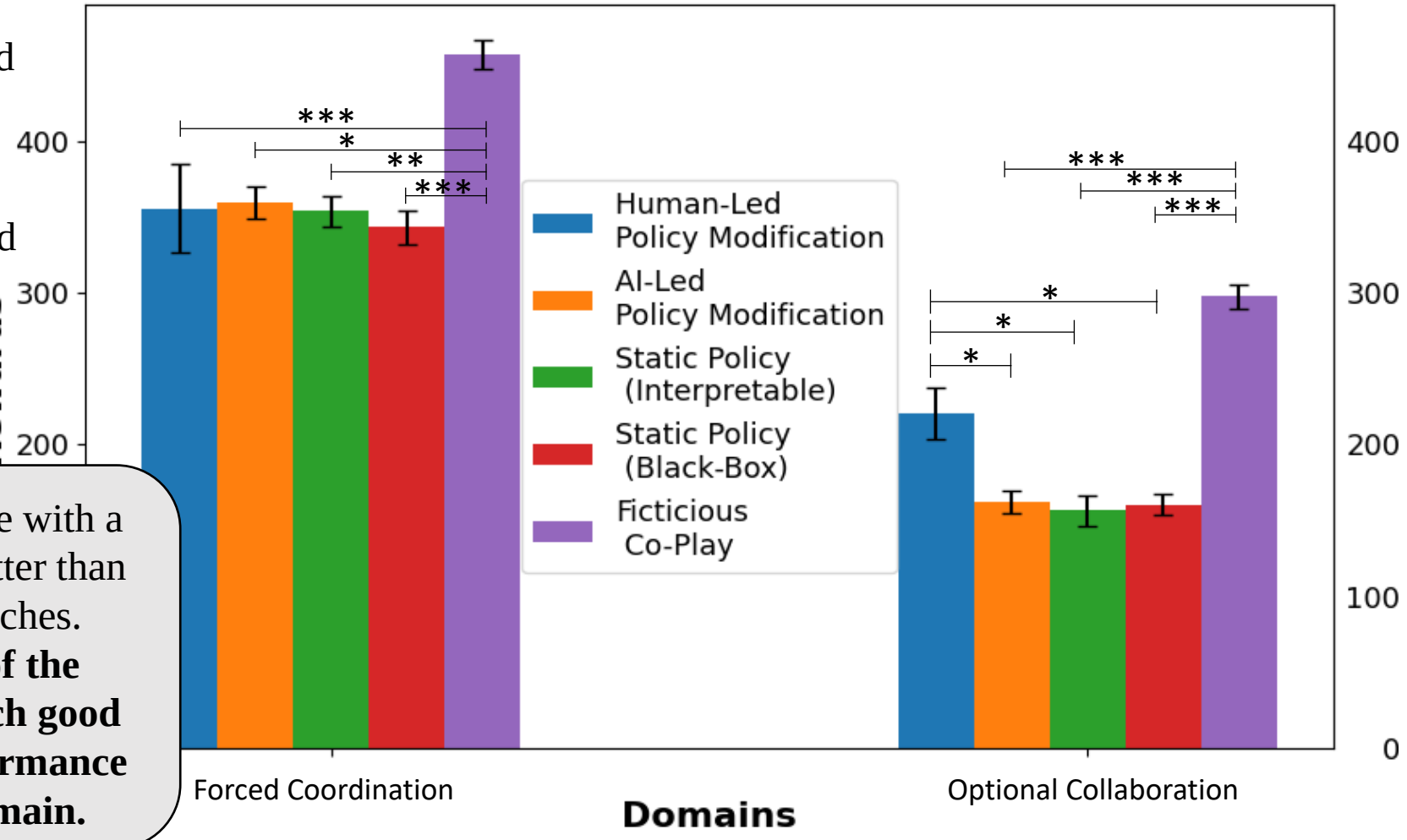
Team Coordination Performance

Collaborative Heuristic
408

Individual Heuristic Reward
306

NN Training Reward
403.16 ± 16.08

IDCT Training Reward
315.22 ± 14.59

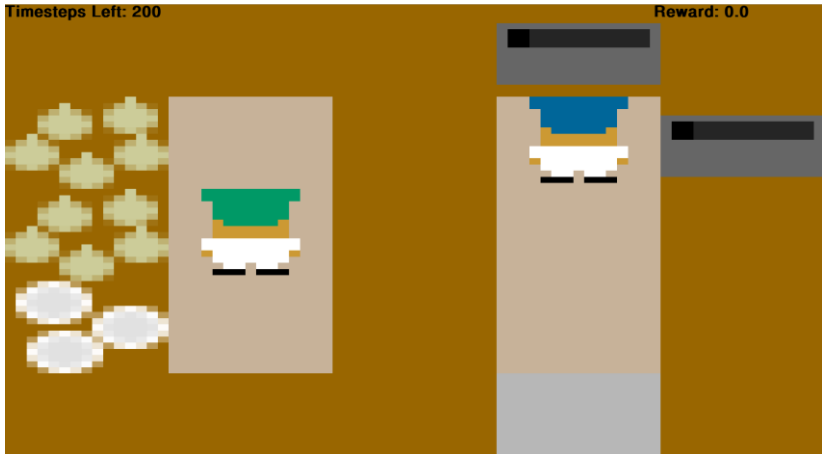


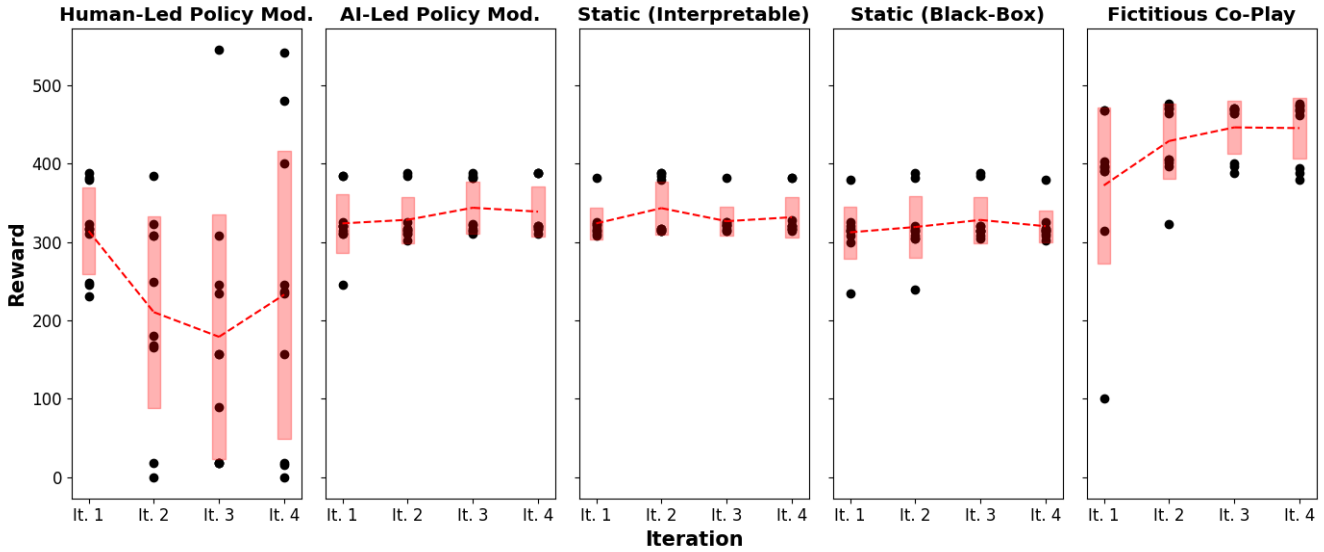
NN Training Reward
295.02 ± 1.86

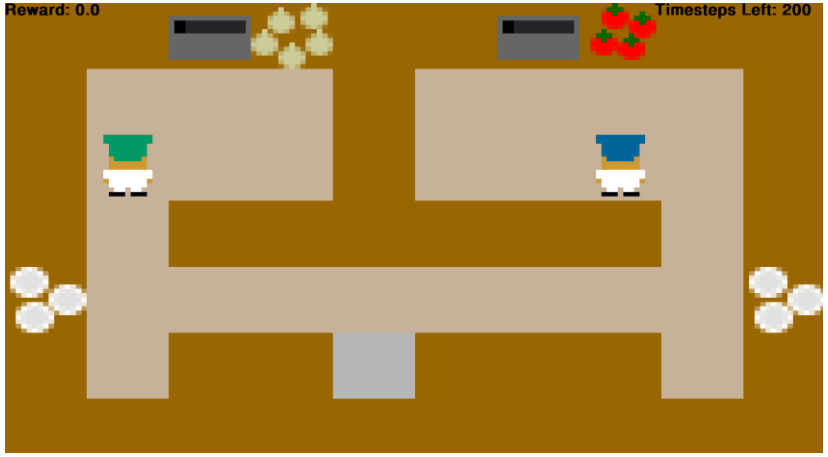
IDCT Training Reward
171.46 ± 18.89

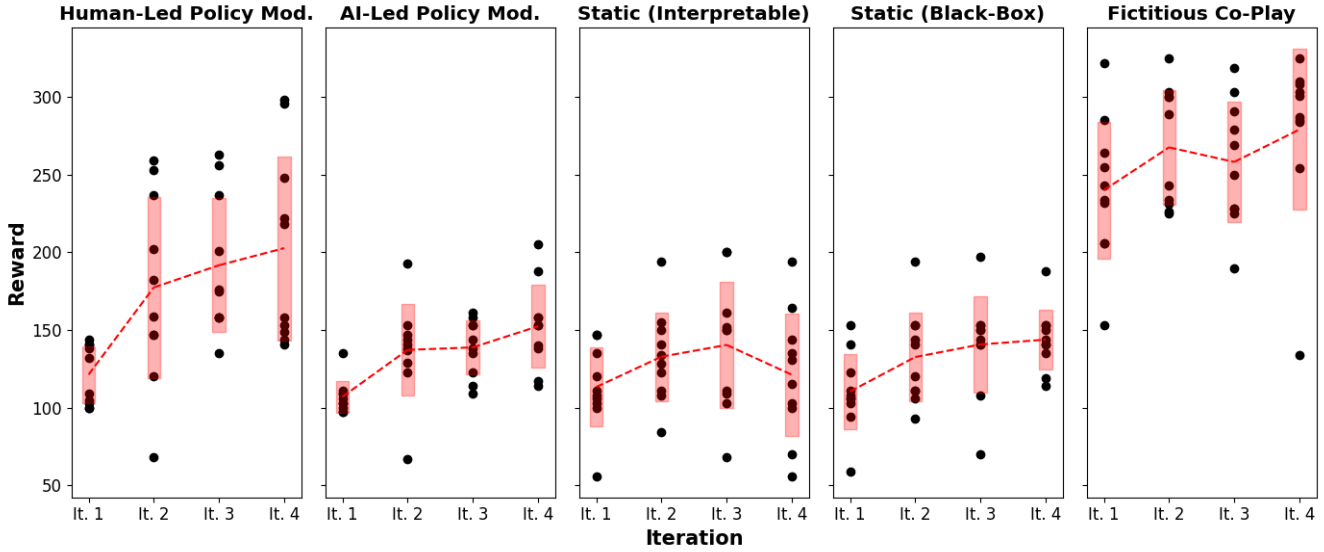
Users can coordinate with a black-box model better than white-box approaches. **However, none of the conditions approach good collaborative performance in the second domain.**

Team Development

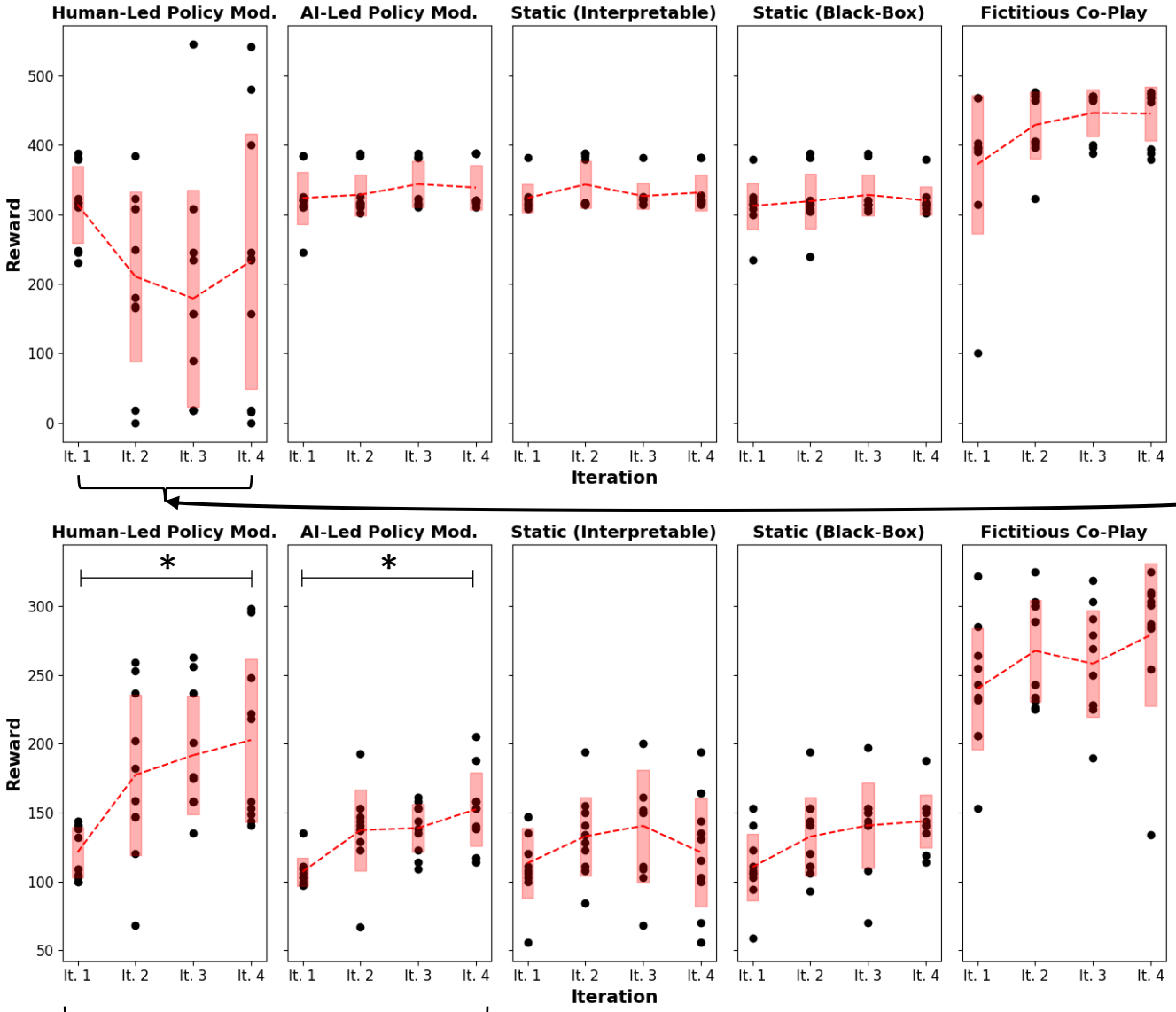


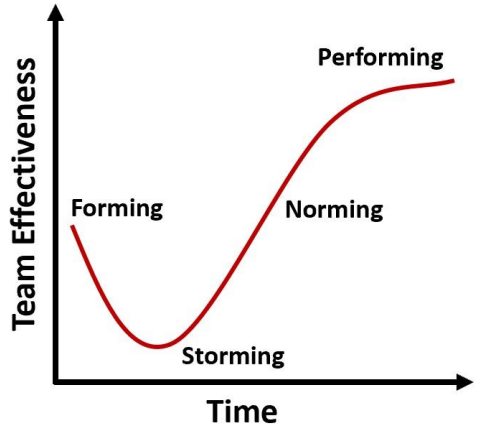






Team Development





Performance first decreases (**forming and storming stages**) and begins to increase (**norming**). In future, we would like to evaluate a larger number of iterations to see if the behavior would continue to trend upward.

We find a significant effect between improvement and familiarity with decision trees (p<0.01)

White-box approaches with policy modification benefited team improvement over repeated play, facilitating the **norming stage** of Tuckman's model

Design Guidelines to Improve Human-Machine Teaming

1. The creation of white-box learning approaches that can produce interpretable collaborative agents that achieve competitive initial performance to that of black-box agents.
2. The design of learning schemes to support the generation of collaborative AI behaviors rather than individual coordination.
3. The creation of mixed-initiative interfaces that enable users, who may vary in ability and experience, to improve team collaboration across and within interactions.
4. The evaluation of teaming in a larger number of interactions.

We have many more results in our paper!

[pdf]

[code]