On the Utility of Learning about Humans for Human-AI Coordination

Micah Carroll, Rohin Shah, Mark Ho,
Tom Griffiths, Sanjit Seshia, Pieter Abbeel, Anca Dragan

Center for Human-Compatible Artificial Intelligence

Berkeley UNIVERSITY OF CALIFORNIA

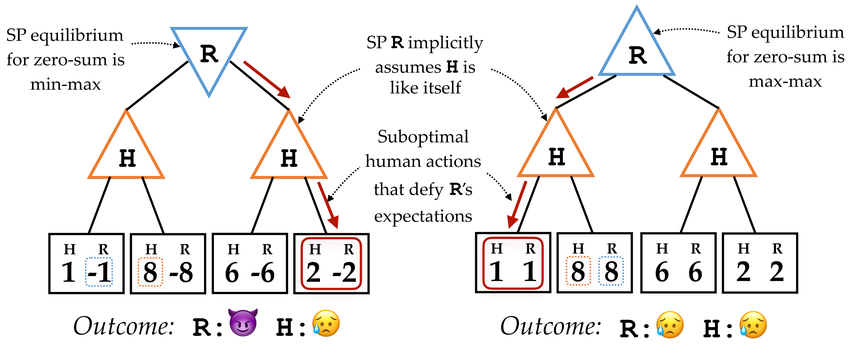
PRINCETON UNIVERSITY

Agents trained with self-play may not coordinate with humans

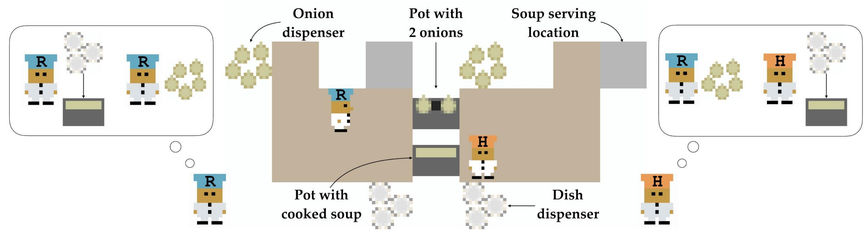
Self play agents face distributional shift when paired with humans. Why are they robust to it in Go, Dota, etc.? Because these are all zero-sum games.

Zero-Sum Games
(e.g. Go, Starcraft, Quake, Dota*)

Cooperative Games
(e.g. Overcooked)



When agents incorrectly expect humans to be optimal, coordination failures result.



Even though **H** is clearly headed for a dish, **R** expects **H** to turn around and get an onion, and so gets a dish itself

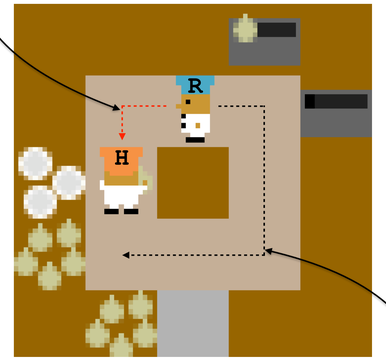
In Overcooked coordination is essential for achieving high performance

***Key Insight**: Assuming that your opponent is optimal is a conservative assumption in zero-sum games but not in cooperative games*

Adaptation to humans

Agents that modeled humans are significantly more adaptive and can take on both leader and follower roles.

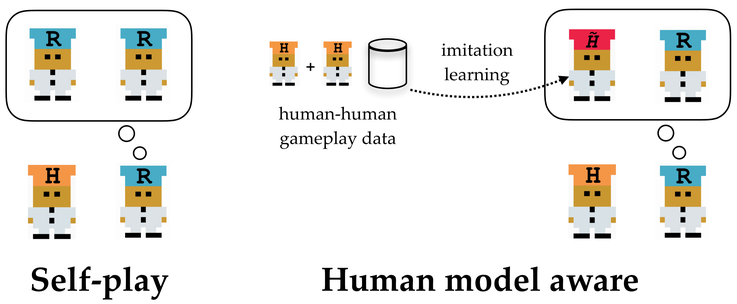
SP would crash into the human, expecting it to turn and go the other way (as itself would do)



Human model aware agents learn that if the human is in the current state, it's better to "give way"

Could data from humans help?

What if we used knowledge of the human to train to expect human suboptimality, and thus suffer from less distributional shift?



We hypothesize that in cooperative settings:

H1. *Self-play agents will perform much more poorly when paired with humans.*

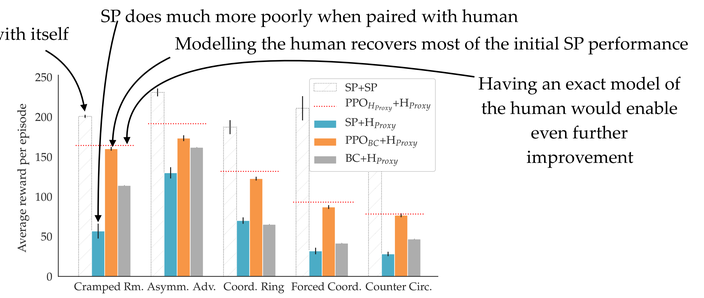
H2. *Using human data or models for training will lead to significant improvements.*

H3. *RL with human models will outperform vanilla imitation learning.*

Yes it helps!

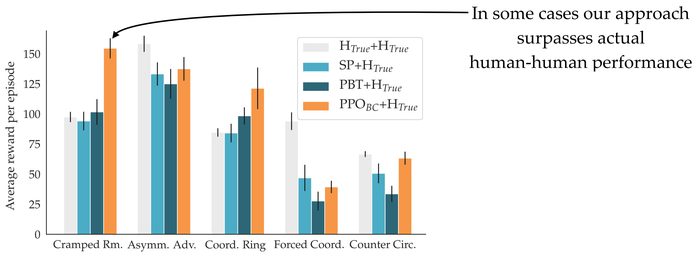
Our human-model-aware approach performs significantly better than self-play

Results with held-out simulated human model



In a user study we find similar results, although less extreme, as humans learn to work around the self-play agent's coordination failures

Results from user study with real humans



Further improvements

Among the issues that we would like to further explore are:

1. *Trying to obtain human models that are more robust to distributional shift* by using inductive biases such as Theory of Mind.

2. *Biasing population-based training towards human policies* by including human models in the agent population.

3. *Adapting to humans at test time,* potentially through meta-learning.

Further Information

http://humancompatible.ai/