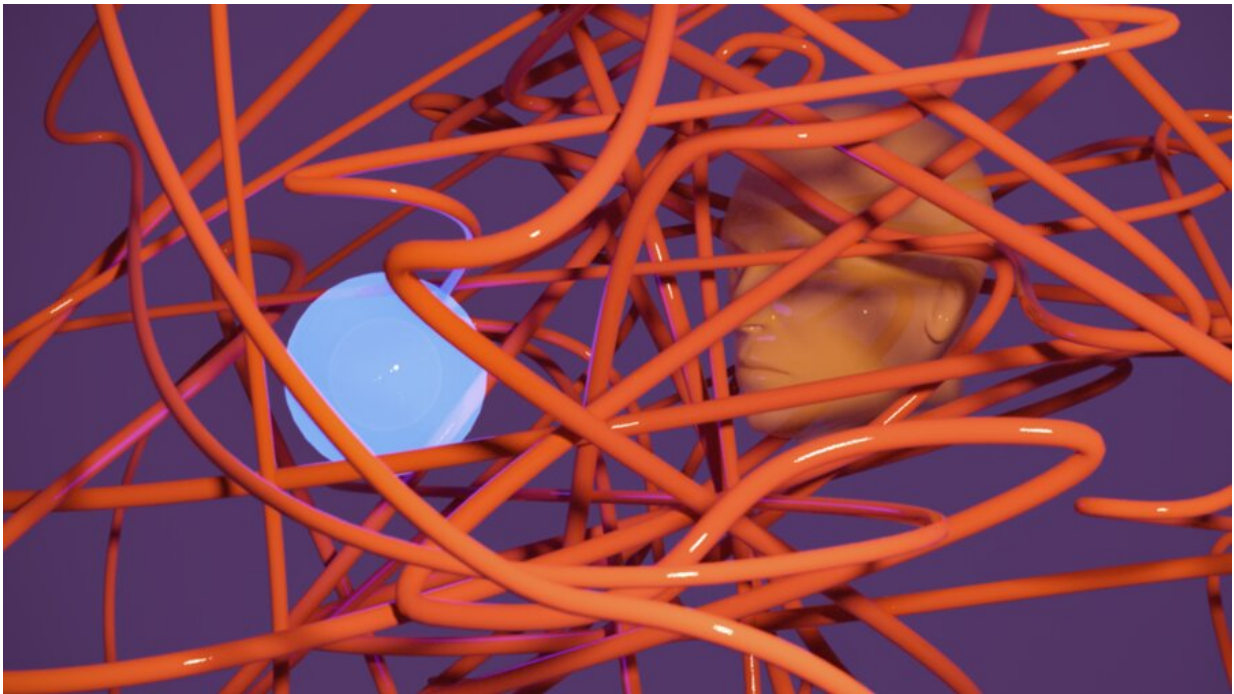# Artificial intelligence is smart, but does it play well with others?

October 4 2021, by Kylie Foy

When playing the cooperative card game Hanabi, humans felt frustrated and confused by the moves of their AI teammate. Credit: Bryan Mastergeorge

When it comes to games such as chess or Go, artificial intelligence (AI) programs have far surpassed the best players in the world. These "superhuman" AIs are unmatched competitors, but perhaps harder than competing against humans is collaborating with them. Can the same technology get along with people?

In a new study, MIT Lincoln Laboratory researchers sought to find out how well humans could play the cooperative card game Hanabi with an advanced AI model trained to excel at playing with teammates it had never met before. In single-blind experiments, participants played two series of the game: One with the AI agent as their teammate, and the other with a rule-based agent, a bot manually programmed to play in a predefined way.

The results surprised the researchers. Not only were the scores no better with the AI teammate than with the rule-based agent, but humans consistently hated playing with their AI teammate. They found it to be unpredictable, unreliable, and untrustworthy, and felt negatively even when the team scored well. A paper detailing this study has been accepted to the 2021 Conference on Neural Information Processing Systems (NeurIPS).

"It really highlights the nuanced distinction between creating AI that performs objectively well and creating AI that is subjectively trusted or preferred," says Ross Allen, co-author of the paper and a researcher in the Artificial Intelligence Technology Group. "It may seem those things are so close that there's not really daylight between them, but this study showed that those are actually two separate problems. We need to work on disentangling those."

Humans hating their AI teammates could be of concern for researchers designing this technology to one day work with humans on real challenges—like defending from missiles or performing complex surgery. This dynamic, called teaming intelligence, is a next frontier in AI research, and it uses a particular kind of AI called reinforcement learning.

A reinforcement learning AI is not told which actions to take, but instead discovers which actions yield the most numerical "reward" by trying out

scenarios again and again. It is this technology that has yielded the superhuman chess and Go players. Unlike rule-based algorithms, these AI aren't programmed to follow "if/then" statements, because the possible outcomes of the human tasks they're slated to tackle, like driving a car, are far too many to code.

"Reinforcement learning is a much more general-purpose way of developing AI. If you can train it to learn how to play the game of chess, that agent won't necessarily go drive a car. But you can use the same algorithms to train a different agent to drive a car, given the right data," Allen says. "The sky's the limit in what it could, in theory, do."

Bad hints, bad plays

Today, researchers are using Hanabi to test the performance of reinforcement learning models developed for collaboration, in much the same way that chess has served as a benchmark for testing competitive AI for decades.

The game of Hanabi is akin to a multiplayer form of Solitaire. Players work together to stack cards of the same suit in order. However, players may not view their own cards, only the cards that their teammates hold. Each player is strictly limited in what they can communicate to their teammates to get them to pick the best card from their own hand to stack next.

The Lincoln Laboratory researchers did not develop either the AI or rule-based agents used in this experiment. Both agents represent the best in their fields for Hanabi performance. In fact, when the AI model was previously paired with an AI teammate it had never played with before, the team achieved the highest-ever score for Hanabi play between two unknown AI agents.

"That was an important result," Allen says. "We thought, if these AI that have never met before can come together and play really well, then we should be able to bring humans that also know how to play very well together with the AI, and they'll also do very well. That's why we thought the AI team would objectively play better, and also why we thought that humans would prefer it, because generally we'll like something better if we do well."

Neither of those expectations came true. Objectively, there was no statistical difference in the scores between the AI and the rule-based agent. Subjectively, all 29 participants reported in surveys a clear preference toward the rule-based teammate. The participants were not informed which agent they were playing with for which games.

"One participant said that they were so stressed out at the bad play from the AI agent that they actually got a headache," says Jaime Pena, a researcher in the AI Technology and Systems Group and an author on the paper. "Another said that they thought the rule-based agent was dumb but workable, whereas the AI agent showed that it understood the rules, but that its moves were not cohesive with what a team looks like. To them, it was giving bad hints, making bad plays."

Inhuman creativity

This perception of AI making "bad plays" links to surprising behavior researchers have observed previously in reinforcement learning work. For example, in 2016, when DeepMind's AlphaGo first defeated one of the world's best Go players, one of the most widely praised moves made by AlphaGo was move 37 in game 2, a move so unusual that human commentators thought it was a mistake. Later analysis revealed that the move was actually extremely well-calculated, and was described as "genius."

Such moves might be praised when an AI opponent performs them, but they're less likely to be celebrated in a team setting. The Lincoln Laboratory researchers found that strange or seemingly illogical moves were the worst offenders in breaking humans' trust in their AI teammate in these closely coupled teams. Such moves not only diminished players' perception of how well they and their AI teammate worked together, but also how much they wanted to work with the AI at all, especially when any potential payoff wasn't immediately obvious.

"There was a lot of commentary about giving up, comments like "I hate working with this thing,'" adds Hosea Siu, also an author of the paper and a researcher in the Control and Autonomous Systems Engineering Group.

Participants who rated themselves as Hanabi experts, which the majority of players in this study did, more often gave up on the AI player. Siu finds this concerning for AI developers, because key users of this technology will likely be domain experts.

"Let's say you train up a super-smart AI guidance assistant for a missile defense scenario. You aren't handing it off to a trainee; you're handing it off to your experts on your ships who have been doing this for 25 years. So, if there is a strong expert bias against it in gaming scenarios, it's likely going to show up in real-world ops," he adds.

Squishy humans

The researchers note that the AI used in this study wasn't developed for human preference. But, that's part of the problem—not many are. Like most collaborative AI models, this model was designed to score as high as possible, and its success has been benchmarked by its objective performance.

If researchers don't focus on the question of subjective human preference, "then we won't create AI that humans actually want to use," Allen says. "It's easier to work on AI that improves a very clean number. It's much harder to work on AI that works in this mushier world of human preferences."

Solving this harder problem is the goal of the MeRLin (Mission-Ready Reinforcement Learning) project, which this experiment was funded under in Lincoln Laboratory's Technology Office, in collaboration with the U.S. Air Force Artificial Intelligence Accelerator and the MIT Department of Electrical Engineering and Computer Science. The project is studying what has prevented collaborative AI technology from leaping out of the game space and into messier reality.

The researchers think that the ability for the AI to explain its actions will engender trust. This will be the focus of their work for the next year.

"You can imagine we rerun the experiment, but after the fact—and this is much easier said than done—the human could ask, 'Why did you do that move, I didn't understand it?' If the AI could provide some insight into what they thought was going to happen based on their actions, then our hypothesis is that humans would say, 'Oh, weird way of thinking about it, but I get it now,' and they'd trust it. Our results would totally change, even though we didn't change the underlying decision-making of the AI," Allen says.

Like a huddle after a game, this kind of exchange is often what helps humans build camaraderie and cooperation as a team.

"Maybe it's also a staffing bias. Most AI teams don't have people who want to work on these squishy humans and their soft problems," Siu adds, laughing. "It's people who want to do math and optimization. And that's the basis, but that's not enough."

Mastering a game such as Hanabi between AI and humans could open up a universe of possibilities for teaming intelligence in the future. But until researchers can close the gap between how well an AI performs and how much a human likes it, the technology may well remain at machine versus human.

More information: Evaluation of Human-AI Teams for Learned and Rule-Based Agents in Hanabi, arXiv:2107.07630v2 [cs.AI] arxiv.org/abs/2107.07630

Provided by Massachusetts Institute of Technology