AI assistant monitors teamwork to promote effective collaboration

August 20 2024, by Alex Shipps

A system developed by MIT CSAIL researchers can oversee a team of both human and AI agents, communicating with team members to align roles and accomplish a common goal. Credit: Alex Shipps/MIT CSAIL

On a research cruise around Hawaii in 2018, Yuening Zhang SM '19,

Ph.D. '24 saw how difficult it was to keep a tight ship. The careful coordination required to map underwater terrain could sometimes lead to a stressful environment for team members, who might have different understandings of which tasks must be completed in spontaneously changing conditions.

During these trips, Zhang considered how a robotic companion could have helped her and her crewmates achieve their goals more efficiently.

Six years later, as a research assistant in the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), Zhang developed what could be considered a missing piece: an AI assistant that communicates with team members to align roles and accomplish a common goal.

In a paper presented at the International Conference on Robotics and Automation (ICRA 2024) and published in the *2024 IEEE International Conference on Robotics and Automation (ICRA)* on August 8, she and her colleagues present a system that can oversee a team of both human and AI agents, intervening when needed to potentially increase teamwork effectiveness in domains like search-and-rescue missions, medical procedures, and strategy video games.

The CSAIL-led group has developed a theory of mind model for AI agents, which represents how humans think and understand each other's possible plan of action when they cooperate on a task.

By observing the actions of its fellow agents, this new team coordinator can infer their plans and their understanding of each other from a prior set of beliefs. When their plans are incompatible, the AI helper intervenes by aligning their beliefs about each other, instructing their actions, as well as asking questions when needed.

For example, when a team of rescue workers is out in the field to triage

victims, they must make decisions based on their beliefs about each other's roles and progress. This type of epistemic planning could be improved by CSAIL's software, which can send messages about what each agent intends to do or has done to ensure task completion and avoid duplicate efforts.

In this instance, the AI helper may intervene to communicate that an agent has already proceeded to a certain room, or that none of the agents are covering a certain area with potential victims.

"Our work takes into account the sentiment that 'I believe that you believe what someone else believes,'" says Zhang, who is now a research scientist at Mobi Systems. "Imagine you're working on a team and you ask yourself, 'What exactly is that person doing? What am I going to do? Does he know what I am about to do?' We model how different team members understand the overarching plan and communicate what they need to accomplish to help complete their team's overall goal."

AI to the rescue

Even with a sophisticated plan, both human and robotic agents will encounter confusion and even make mistakes if their roles are unclear. This plight looms especially large in search-and-rescue missions, where the objective may be to locate someone in danger despite limited time and a vast area to scan.

Thankfully, communication technology augmented with the new robotic assistant could potentially notify the search parties about what each group is doing and where they're looking. In turn, the agents could navigate their terrain more efficiently.

This type of task organization could aid in other high-stakes scenarios such as surgeries. In these cases, the nurse first needs to bring the patient

to the operation room, then the anesthesiologist puts the patient to sleep before the surgeons begin the operation.

Throughout the operation, the team must continuously monitor the patient's condition while dynamically responding to the actions of each colleague. To ensure that each activity within the procedure remains well-organized, the AI team coordinator could oversee and intervene if confusion about any of these tasks arises.

Effective teamwork is also integral to video games such as "Valorant," where players collaboratively coordinate who needs to attack and defend against another team online. In these scenarios, an AI assistant could pop up on the screen to alert individual users about where they've misinterpreted which tasks they need to complete.

Before she led the development of this model, Zhang designed EPike, a [computational model](link) that can act as a team member. In a 3D simulation program, this algorithm controlled a robotic agent that needed to match a container to the drink chosen by the human. As rational and sophisticated as they may be, cases arise where these AI-simulated bots are limited by their misconceptions about their human partners or the task.

The new AI coordinator can correct the agents' beliefs when needed to resolve potential problems, and it consistently intervened in this instance. The system sent messages to the robot about the human's true intentions to ensure it matched the container correctly.

"In our work on human-robot collaboration, we've been both humbled and inspired over the years by how fluid human partners can be," says Brian C. Williams, MIT professor of aeronautics and astronautics, CSAIL member, and senior author on the study.

"Just look at a young couple with kids, who work together to get their kids breakfast and off to school. If one parent sees their partner serving breakfast and still in their bathrobe, the parent knows to shower quickly and shuffle the kids off to school, without the need to say a word. Good partners are well in tune with the beliefs and goals of each other, and our work on epistemic planning strives to capture this style of reasoning."

The researchers' method incorporates probabilistic reasoning with recursive mental modeling of the agents, allowing the AI assistant to make risk-bounded decisions. In addition, they focused on modeling agents' understanding of plans and actions, which could complement previous work on modeling beliefs about the current world or environment.

The AI assistant currently infers agents' beliefs based on a given prior set of possible beliefs, but the MIT group envisions applying machine learning techniques to generate new hypotheses on the fly. To apply this counterpart to real-life tasks, they also aim to consider richer plan representations in their work and reduce computation costs further.

Provided by Massachusetts Institute of Technology