Human-AI Collaborations for Controlled Tasking Use Cases

Michael Atighetchi, Sedona Thomas, Regan Broderick-Sander, Julia Hiebel
RTX BBN Technologies

matighet@bbn.com, sthomas@bbn.com, regan.broderick-sander@bbn.com, julia.hiebel@bbn.com

Abstract

Many Human-AI collaborations are based on task environments in which tasks on controllable assets need to be taken (e.g., to change a route of a truck or to send an informational message to the truck driver) in response to results of monitoring tasks (e.g., changes in customer needs, changes in weather). Human-AI teams need to collaborate effectively and efficiently to generate, assess, and adjust Courses of Actions (COAs) in those environments. RTX BBN Technologies (BBN) has developed multiple AI agents, covering different specific use cases, and is currently deploying Human-AI teams into operational environments. In our work, the most relevant use cases involve an AI agent advising a human on executing complex tasks and human and AI agents handling subtasks usually done by one person. The workshop presentation will focus on overall context, our successes with Human-AI teams in recent years, and specific insights from our work with Human-AI teams performing joint activities in controlled tasking environments.

Problem Domain and Overall Objective of Human-AI Collaboration

An abstracted version of the problem domain can be described as the management of activities within the operational environment based on the commands, direction, and guidance given by an appropriate authority. While our work involves specific use cases (Atighetchi et al. 2020), this discussion abstracts tasks to general types of monitoring the environment, redirecting assets, and informing assets, where assets could be delivery trucks or commercial aircraft.

While specific objectives vary across specific use cases, we assert that each use case has a means to measure performance. In the truck delivery context, those metrics might include time-to-deliver goods and percentage of goods delivered on time. Such metrics enable measuring the performance of Human-AI teams over Human-only teams.

Human and AI Team Members, Including Abilities and Expertise

Controlled tasking environments of a certain scale decompose tasks into a set number of roles, e.g., overall coordinators, monitoring support, and weather specialists. Similarly, AI team members decompose into functional units, e.g., a microservice for weather monitoring, a microservice for generating COAs, and a microservice for performing path prediction. AI team members interact with each other and with humans, providing a rich environment for reinforced feedback and targeted collaboration. User input, together with contextual information, enables these AI agents to improve performance over time and learn to customize interactions to specific human operators.

Roles and Responsibilities of Team Members for Component Tasks

We have seen component tasks which can categorized into the following three sets.

AI-only tasks involving data intensive and/or repetitive activities, e.g., predicting paths and turning speech to text.

Tasks jointly executed by humans and AI when decision making between the two is closely aligned. For instance, an AI team member might suggest COAs to a human who is likewise mentally forming expected COAs. These tasks follow the model of an AI team member providing decision support to a human as a means of validating or surprising human expectation and decision making.

Tasks only executed by humans involving decisions and actions with larger operational impact. For instance, while changing the route of a single delivery track might be performed by AI agents without human intervention, changing routes of multiple tracks will only be allowed if the AI team member has built a level of trust with humans (as certain this could significantly impact a large amount of customer dissatisfaction if not done right), and stopping all track deliveries for a larger area will always be a human-only task, as it could put the company out of business.

Objectives, Expectations, and Interactions among Team Members

In controlled tasking environments, human operators generally expect AI team members to be able to act within seconds, avoid thrashing, continuously update plans and predictions based on evolving environmental conditions every second, and display results in an easy to intuit, non-cluttering, and fully integrated User Interface.

Furthermore, we have seen interactions and expectations amongst team members evolve over time. For our use cases, initial expectations for the COA AI agent were to tell it the high-level effect that should be achieved (e.g., reroute target destination to X, pick up more goods at Y, avoid area Z). As human team members' confidence in the COA AI agent grew, operators became willing to relinquish this level of decision authority to AI agents, allowing them to autonomously select the right effects.

While human team members are willing to take COA suggestions from an AI team member, there continues to be real hesitancy in allowing AI team member to execute top-ranking COAs autonomously. Autonomous COA execution is established for very limited conditions, but lack of trust in AI agents is frequently cited as the main roadblock for further adoption.

Incorporation of Features Like Transparency, Trust, and Dependability

Our work on controlled tasking has included aspects to increase transparency and trust. For monitoring activities, the ability to drill down into individual features has been critical. Humans care more about the type and quantity of features that indicate a condition rather than numeric scores, even when human team members can configure score mappings for each feature.

For COAs, we found that the ability to explain why certain assets were not included in any COAs was critical to answer questions posed by operators, e.g., "Why was delivery truck X was not used?". In addition, for showing tradeoffs between different COAs in metrics space, we found that a small number of high-level metrics (timeliness, probability of success, collateral damage, opportunity cost) were sufficient to covey important tradeoffs.

Responses to Benefit-risk Tradeoffs in AI-Human Collaboration

The BBN COA generation AI generates sets of consistent sets of COAs, where each consistent set provides COAs that impose specific effects on a specific set of targets. In the truck domain, the effect would be to supply certain target locations with certain goods. When generating the sets of consistent sets, our algorithm identifies tradeoffs across metrics, enabling human operators to stay in control of the options space without becoming overwhelmed by options.

Another aspect of our COA generation AI capability is that humans can specify preference models which outline their preferred operating point in a multi-dimensional metrics space. This means that human team members are able to, for example, express their preferred tradeoff between timeliness, probability of successful delivery, and fuel costs. The AI is subsequently able to provide options which have been optimized against the provided preference model.

Measures of Task Performance for Overall Team and Members

As mentioned earlier, specific controlled tasking environments have specific operational requirements and criteria that need to be met for tasks to succeed. For monitoring activities, the main metrics include the ability to generate alerts, provide meaningful contextual information about assets and environmental conditions.

Similarly, the performance of COA generation is measured by the ability to support a large and complex set of reconfigurations. In the trucking domain, this involves handling different types of trucks, their fuel rates and capacities, and to provide COAs that maximize the probability of success subject to several other constraints.

Conclusion

BBN's work in specific applications of the controlled tasking domain provides a compelling use case for human-AI collaboration, with real-life application to national security and homeland defense. Participation in this workshop enables us to provide high-level lessons learned to the community and explore new directions on closing gaps between AI and human team members. BBN's participation will bring an applied yet generalized perspective, enabling discussions at an unrestricted level with the larger community.

References

Atighetchi, M.; Broderick-Sander, R.; Cook, I.; Last, D. 2023. Risk Assessment & Course of Action Generation for Homeland Air Defense. Paper presented at the Restricted Track of IEEE Military Network Conference