Inferring Player and Team Models in a Minecraft Search-and-Rescue Task

David V. Pynadath, Nikolos Gurney
Institute for Creative Technologies
University of Southern California
Los Angeles CA
{pynadath, gurney}@ict.usc.edu

Stacy C. Marsella
Northeastern University
Boston, MA
s.marsella@northeastern.edu

Hala Mostafa
BitSight
Boston, MA
hala.mostafa@bitsight.com

Pedro Sequeira
SRI International
Menlo Park, CA
pedro.sequeira@sri.com

Volkan Ustun
Institute for Creative Technologies
University of Southern California
Los Angeles CA
ustun@ict.usc.edu

Peggy Wu
Raytheon Technologies Research Center
East Hartford, CT
peggy.wu@rtx.com

ABSTRACT

While autonomous agents offer the promise of automated assistance for human teamwork, the usefulness of such assistance is limited by an agent's ability to understand the people it wants to help. No matter how well defined and tightly scoped a task is, people will always bring their diverse preferences, experiences, biases, etc. to bear even when acting as a team. An agent that can neither represent the heterogeneity across the team it will have to help, nor infer the subjective frame of reference of the individual team members, is doomed to apply a one-size-fits-all policy of assistance that is unlikely to maximize task performance. In addition to the heterogeneity of team members, there is also heterogeneity across the tasks that a team might be asked to perform. For our agent to accurately assess and assist diverse people performing diverse joint tasks, it must be an *artificial social intelligence* (ASI) whose inference and interventions can generalize across people and tasks. In this work, we present a methodology for capturing the diversity across different teams of people within an agent's beliefs and reasoning about interventions to assist them. The agent here does not directly observe the task environment, but instead relies on input from *analytic components* that process information from that environment and output only teamwork-relevant measures. Our agent models these teamwork variables and updates its beliefs over them using a Bayesian Theory of Mind, by applying Partially Observable Markov Decision Processes in a recursive manner to assess the state of the team it is currently observing and to choose interventions to best assist them. We present this methodology using a search-and-rescue task in Minecraft and propose an evaluation scheme to be deployed upon completion of data collection in that task.

ABOUT THE AUTHORS

David V. Pynadath is the Director for Social Simulation Research at the USC Institute for Creative Technologies and a Research Assistant Professor in the USC Computer Science Department. He has published papers on social simulation, multiagent systems, teamwork, plan recognition, and adjustable autonomy. He is the co-creator and maintainer of PsychSim, a multiagent social simulation framework that has been used in interactive simulations for teaching urban stabilization operations, cross-cultural negotiation, and avoiding risky behavior. Dr. Pynadath's work on PsychSim is a key component of his long-term research into applying decision-theoretic multiagent methods to models of behavior. He has developed multiagent systems for applications in social simulation, virtual training environments, human-robot interaction, automated personal assistants, and UAV coordination. He has used such systems to create models of human decision-making in scenarios including ethnic conflict, traffic, classroom violence, negotiation, and disaster response.

Nikolos Gurney a postdoctoral research associate at the USC Institute for Creative Technologies applying behavioral science insights to the development of AI and using deep learning methods to uncover new knowledge about human behavior. Prior to ICT, he completed a PhD in Behavioral Decision Research at Carnegie Mellon University.

Stacy C. Marsella is a Professor at Northeastern University in the Khoury College of Computer Sciences with a joint appointment in the Department of Psychology. His research is in the computational modeling of human cognition, emotion and social behavior, both as a basic research method in the study of human behavior as well as for use in a range of applications. His work has been applied to the modeling of human behavior for large scale social simulations, realization of effective human-AI teamwork as well as design of virtual humans, software entities that look human and can interact with humans in virtual environments using verbal and nonverbal behavior. He is the co-creator of the PsychSim multi-agent social simulation framework.

Hala Mostafa leads the Security Analytics Team at Bitsight, a cybersecurity ratings leader that analyzes vast amounts of data to evaluate the cybersecurity posture of companies and government agencies. Before that, Mostafa spent 10 years doing basic and applied AI/ML research in the defense and aerospace industries. Mostafa holds a PhD in Computer Science from the University of Massachusetts, Amherst.

Pedro Sequeira is a Senior Computer Scientist working at SRI International in the Artificial Intelligence Center (AIC). Dr. Sequeira mainly has expertise in the fields of Artificial Intelligence (AI) / Machine Learning (ML) and in particular Reinforcement Learning (RL), involving the creation of autonomous and adaptive systems that learn and reason under uncertainty. His research interests include the use of ML to better understand how humans learn and make decisions in complex tasks and helping humans understand the behavior of artificial agents through explainability, with the goal of improving Human-AI collaboration. Dr. Sequeira has been the recipient of several awards for his research and is the author of more than 30 peer-reviewed scientific publications.

Volkan Ustun is the Associate Director of the Cognitive Architecture group at the USC Institute for Creative Technologies. His general research interests are multi-agent systems, cognitive architectures, computational cognitive models, and simulation.

Peggy Wu has 20 years of experience combining cognitive psychology with Artificial Intelligence to advance Human Computer Interactions, Social Computing, Human Machine Trust, & Ethical AI. She applies advanced technologies such as Virtual/Augmented Reality to improve Human+Machine performance on Earth and in Space. Peggy is the recipient of numerous grants from the DoD, NASA, and DoE, and serves in a leadership capacity for a matrix organization of 200+ team members. She is an inventor with 20 patents (14 pending) and over 70 peer reviewed publications.

Inferring Player and Team Models in a Minecraft Search-and-Rescue Task

David V. Pynadath, Nikolos Gurney
Institute for Creative Technologies
University of Southern California
Los Angeles CA
{pynadath, gurney}@ict.usc.edu

Stacy C. Marsella
Northeastern University
Boston, MA
s.marsella@northeastern.edu

Hala Mostafa
BitSight
Boston, MA
hala.mostafa@bitsight.com

Pedro Sequeira
SRI International
Menlo Park, CA
pedro.sequeira@sri.com

Volkan Ustun
Institute for Creative Technologies
University of Southern California
Los Angeles CA
ustun@ict.usc.edu

Peggy Wu
Raytheon Technologies Research Center
East Hartford, CT
peggy.wu@rtx.com

INTRODUCTION

While autonomous agents offer the promise of automated assessment and assistance for human teamwork, their usefulness is limited by an agent's ability to understand the people it wants to help. No matter how well defined and tightly scoped a task is, people will always bring their diverse preferences, experiences, biases, etc. to bear even when acting as a team. An agent that can neither represent the heterogeneity across the team it will have to help, nor infer the subjective frame of reference of the individual team members, is doomed to apply a one-size-fits-all approach that is unlikely to maximize accuracy or performance.

In addition to the heterogeneity of team members, there is also heterogeneity across the tasks that a team might be asked to perform. A good agent should be able to adapt to changes in the task environment with minimal modification, if not none at all. In particular, while it may need to incorporate new task knowledge, its *teamwork* reasoning should not change.

For our agent to accurately assess and assist diverse people performing diverse joint tasks, it must be an *artificial social intelligence* (ASI) whose inference and interventions can generalize across people and tasks. More precisely, it must be able to abstract away the task-dependent aspects and reason at a team level. It must then adapt its team-level reasoning to the dynamic state of the specific team and the specific team members that it is currently assisting.

In this work, we present a methodology for capturing the diversity across different teams of people within an agent's beliefs and reasoning about interventions to assist them. The agent here does not directly observe the task environment, but instead relies on input from *analytic components* (ACs) that process information from that environment and output only teamwork-relevant measures. Our agent models these teamwork variables and updates its beliefs over them using a Bayesian Theory of Mind, by applying Partially Observable Markov Decision Processes in a recursive manner to assess the state of the team it is currently observing and to choose interventions to best assist them.

We present this methodology using a search-and-rescue task in Minecraft. We describe the analytic components (developed by other research teams) and how they are incorporated into a model of the underlying team processes. We present the candidate interventions and the selection mechanism used by our agent. Finally, we propose an evaluation scheme to be deployed upon completion of data collection in that task.

BACKGROUND

DARPA's Artificial Social Intelligence for Successful Teams (ASIST) program aims to develop artificial social intelligence (ASI) with human-like social cognition abilities. When humans interact, they maintain and rely on models of each other's belief states, which is frequently referred to as theory of mind (ToM) (brief survey in Gurney et al, 2021). Although there is no consensus on the cognitive processes that constitute human ToM, there is evidence that ToM is built up from observed actions, contextual hints, prior personal experiences, and beliefs. It is also believed that in team settings humans coalesce their individual beliefs, partially through ToM, to create shared mental models of tasks, goals, etc. The ASIST program posits that these same skills, ToM and shared mental models, will prove crucial elements of ASI.

ASIST brings together researchers from a broad array of backgrounds: artificial intelligence, computer science, cognitive science, management science, and psychology to name a few. These researchers work on a variety of different aspects of developing ASI. A handful of teams, including our own, are developing agents capable of social skills similar to humans. Other teams, however, are developing a testbed environment for evaluating the ASI or creating analytical components (ACs) that provide the ASIs with data about their human counterparts.

MINECRAFT TESTBED

One of the biggest challenges to developing robust ASI is the evaluation of its capabilities. This includes obviously necessary capabilities, like modeling the belief states of humans, as well as less obvious capabilities, like being able to model how the belief states of humans change in response to subtle task differences. In situ interaction is too complex and nuanced to facilitate the incremental work needed during early development of ASI. Thus, the ASIST program includes a highly capable testbed (Freeman et al., 2021; Corral et al., 2021), which is implemented in a game-based (Minecraft) environment that facilitates distributed teaming tasks (Johnson et al., 2016).

The first three rounds of experimentation in the ASIST program focused on an urban search and rescue (USAR) task that involves clearing and avoiding hazards while rescuing victims of a disaster. The first round did not involve teams or the ASI. Solo participants simply explored a Minecraft map fashioned in the style of a building looking for and triaging victims. This provided initial data for developing theory around how people engage with the task, hypotheses about how they may team, and generating ideas of what an ASI may contribute to the teams.

The second round of experimentation brought distributed teams of three participants together in an expanded Minecraft map to accomplish a similar mission, again without the presence of ASI. In addition to working as a team to save victims, participants could select different tools to use during the mission. The tools decayed with use and participants could form strategies and plan around which tools to use during different stages of the USAR mission.

The most recent experiment introduces an ASI as a team member in a task that is an evolution of the second round USAR task. Each team is equipped with one of the available ASIs from the program, and the ASI has access to information about each person gathered during an intake survey, output from the ACs that process mission data in real time, and the actual mission data, made available via a message bus. The ASI are not embodied team members, but they are able to communicate, via chat, with the human team members (who can communicate amongst themselves through a shared audio channel). They do so with the explicit goals of improving team processes and facilitating cooperation during the course of the mission.

ANALYTICAL COMPONENTS

ACs provide the ASIs with signals for interpretation. If the ASI is the brain, then the ACs are its myriad senses and sensors that allow it to experience teams of humans. While the ASIs also have access to the detailed state of the task environment from the testbed, we made a conscious decision to not use this task-specific information and rely instead on the team-level abstractions generated by the ACs. Just as there are multiple teams developing different ASIs in the ASIST program, there are six teams using cutting-edge social science to inform ACs that support those ASIs. At present, our ASI makes use of these analytical components, each providing a unique interpretation of the datastream coming out of the testbed. The following is the list of ACs and the subset of their variables that our agent uses in its reasoning:

- Carnegie-Mellon University's (CMU's) Team Effectiveness Diagnostic AC (TED-AC) uses mission data to generate a dynamic stream of four team processes measures (collective effort, appropriate skill use, appropriate use of strategies, and communications) based on Hackman's team coaching framework (Hackman & Wageman, 2005). These measures are hypothesized to predict collective intelligence, including a team's ability to perform across various tasks types.
    - *comms_equity*: the distribution of voice communication across the team (e.g., is one member dominating the conversation)
    - *comms_total_words*: the overall volume of the team's voice communication
    - *inaction_stand_s*: the number of seconds with no activity by the team during the last 30s
    - *process_effort_agg*: the overall task effort spent by the team
- CMU's Background of Experience, Affect, and Resources Diagnostic AC (BEARD-AC) relies on the intake survey, along with participant training performance data, to generate a profile for each team member and the team as a whole along each of the following dimensions: *anger*, *anxiety*, *gaming experience*, *Reading the Mind in the Eyes (RMIE)* (Baron-Cohen, 2001), *Santa Barbara Sense of Direction (SBSOD) scale* (Hegarty et al., 2002).
- Cornell's Trust AC provides measures of players' trust in their teammates, including calculated measures of the players' compliance behaviors (which is widely considered as a proxy for their trusting behaviors). It records the number of *open requests* that exist between each pair of players and the *compliance* of the requestees in each pair in satisfying those requests.
- The Gallup - Emergent Leadership Prediction (GELP) AC provides a measure of emergent leadership for each human team member. These measures are constituted of eight components derived from audio data, NLP, task competency scores, the intake survey data, etc. (Maese et al., 2021). The emergent leadership measure is not static: as more data become available, confidence in a score for a given person changes. These scores do not identify whether a person is potentially a good/bad leader, rather if they are likely to be viewed by their teammates as the understood leader. We use only the overall *Leadership* score, not the component measures computed by this AC.
- The Florida Institute for Human and Machine Cognition (IHMC) Joint Activity AC uses a graph representation of task dependencies to track the team members' activity status and interdependency. It reports on changes in such status, including when a new activity is discovered by a team member, when a team member begins working on an activity, and when such activities are completed. The AC provides status for each possible activity in terms of whether it has been *discovered* and, if so, is the participant *preparing* to address or actually *addressing* it or whether it has been *completed*.
- The Rutgers Utility Agent AC provides a set of measures related to the needs of the team members. This includes information about whether a team member is in danger and if other team members are aware of such situations, how well the team communicates about tasks such as victim triage, and whether team members make mistakes and if other team members recognize that (such as one person moving a victim for triage, but placing the victim in the wrong location). In this case, we make use of the *wait_time* variable to identify potential problem situations.
- University of Central Florida's (UCF's) Player Profiler AC makes predictions of each team member's 'potential' for teamwork and provides ASI with features of each team member that may be predictive of performance related behaviors (Bendell et al., 2021). This potential measure consists of two main parts, *team-potential-category* and *task-potential-category*. The former category represents the categorization of the player as high or low in potential to successfully maintain awareness of their teammates and progress as well as to coordinate activities and resources effectively and efficiently. The latter categorizes team members as high or low in potential to successfully complete mission related actions effectively and efficiently.

TEAM PROCESSES

There is a large body of work analyzing human teamwork and identifying the key variables that distinguish and drive team coordination and performance. In this initial work, we use the team process variables (Marks et al., 2001) whose measures were recently validated (Mathieu et al., 2020). Not all of these variables are relevant in the USAR task environment, so we work with the following subset of variables, as described in earlier publications (Mathieu et al., 2020):

- **Affect management:** "activities that foster emotional balance, togetherness, and effective coping with stressful demands and frustration."
- **Coordination:** "the process of synchronizing or aligning the members' actions."
- **Motivating and confidence building:** "activities that develop and maintain members' motivation and confidence while working toward team goals."
- **Systems monitoring:** "activities such as tracking team resources (e.g., money) and factors in the team environment (e.g., inventories) to ensure that the team has what it needs to accomplish its goals and objectives".
- **Team monitoring:** "members assisting others in the performance of their tasks (by providing feedback or coaching or assisting with the task itself)".

None of these processes are directly observable, nor are the measures developed for them implemented in the program testbed. Our goal here is to model them as hidden variables that our ASI seeks to improve through its interventions. Furthermore, although our ASI cannot directly observe these variables while interacting with the team, there are certain AC variables that provide evidence of the status of these processes, as described in a later section.

PSYCHSIM

Our Agent with Theory Of Mind for Intelligent Collaboration (ATOMIC) ASI is built within PsychSim. PsychSim provides reusable AI technology for generating multiagent systems capable of populating game environments (Pynadath & Marsella, 2005). PsychSim represents individuals and groups as autonomous agents that integrate two AI technologies: recursive models (Gmytrasiewicz & Durfee, 1995) and decision-theoretic reasoning (Kaelbling et al., 1998). Recursive modeling gives agents a Theory of Mind, to form complex attributions about others, incorporate such beliefs into their own behavior, and enrich the explanations provided to the user. In addition, the agents employ boundedly rational, decision-theoretic reasoning to quantitatively assess risk/reward tradeoffs. Thus, these agents represent a decision-making model that generates behavior by reasoning from a declarative representation of their goals and beliefs.

PsychSim was initially developed in conjunction with the Office of the Assistant Secretary of Defense/Special Operations-Low Intensity Conflict (OASD/SOLIC) and the US Army as an exploratory simulation to determine the impact (especially second- and third-order effects) of an influence campaign on a target population (Marsella et al., 2004). PsychSim has since been used to drive the behavior of non-player characters (NPCs) in a variety of serious games. It was used to build entities to populate the DARPA-funded Tactical Language Training System (TLTS), an interactive narrative environment in which students could practice their language and culture skills in a simulated foreign city (Si et al. 2005). We also used PsychSim's mental models and quantitative decision-theoretic reasoning to model a spectrum of negotiation styles within the ELECT BiLAT training system (Kim et al, 2009) (funded by the US Army RDECOM). In the negotiation phase of the training environment, the trainee negotiates with a PsychSim agent that has its own goals and possible negotiation moves (e.g., making offers to the trainee, requesting concessions from the trainee, and accepting/rejecting the overall agreement).

USC ICT's UrbanSim simulation-based training game (McAlinden et al., 2014), funded by the US Army RDECOM, used PsychSim agents to generate behavioral responses to an urban-stabilization operation. UrbanSim allows a trainee to take on the role of a battalion or brigade commander who is attempting to maintain stability, fight insurgency and crime, reconstruct the civic infrastructure and prepare for transition in a fictional megacity, populated by multiple individuals and groups, across diverse social and cultural groups.

In each of these game environments, the PsychSim agents respond to the dynamic situation by applying decision-theoretic AI algorithms to update their beliefs about the state of the world (and of the people in it) and to choose behaviors that best achieve their long-term desires. In prior work, we have used PsychSim to simulate team training in dynamic and heterogeneous environments (Johnson et al., 2019). In fact, PsychSim grew out of a formalism for purely collaborative settings (Pynadath & Tambe, 2002) that we extended to account for people with differing, even adversarial, goals. At the same time, we have used PsychSim to represent a variety of psycho-social phenomena in a domain-independent manner We therefore see PsychSim as a natural foundation for providing domain-independent algorithms for reasoning about teamwork

OUR ASI MODEL

Our goal in building our ATOMIC ASI is to provide ToM-type reasoning at purely the teamwork level, abstracting away any task-specific constructs. We also seek to exploit PsychSim's use of influence diagrams (Howard & Matheson, 2005) to represent the dependency structure among the variables and interventions used by the agent in its decision-making. This graphical representation supports the automatic generation of a potentially human-understandable form to promote transparency and inspection.

This explicit representation of the dependency structure provides a foundation for evaluating specific hypotheses about team-level measures (i.e., from ACs) and underlying (i.e., not directly observable) team process variables. It does so by using a hypothesized structure as an explicit, generative model of team- and individual-level decision-making and behavior, which the ASI then uses recursively.

INTERVENTIONS

We are manually constructing a dynamic influence diagram model that captures various hypothesized dependencies among team-level variables. These include correlations between team-process characteristics and AC-provided profiles of individual players and the team as a whole (e.g., a participant who scores high on CMU's BEARD anxiety scale would contribute to a lower expected affect management capability of the team). These also include effects of team-process characteristics on observed behaviors (e.g., a team with good coordination would have a higher likelihood of advancing their joint activities, as monitored by IHMC's AC). Finally, these include effects of our ASI's candidate interventions on both team-process and AC variables (e.g., cheerleading is likely to increase the team's motivating process variable and CMU TED's measure of collective effort).

The dependency structure and the weights on the links within it are being selected based on background knowledge and significant guesswork. However, the basic capability established here provides a foundation for evaluating the hypothesized model and refining it for future studies as more data come in.

The agent considers the following candidate interventions:

1. **Reflection:** Between trials, the agent prompts the team with a reminder about a situation in the first trial where a player was stuck on a threat plate for an inordinate amount of time.
2. **Cheerleading:** The agent congratulates a player on successfully achieving a goal (more specifically, moving a victim to a triage area).
3. **Report drop:** The agent reports on a noteworthy lack of performance by one player (more specifically, failure to respond to outstanding requests by a teammate).
4. **Notify phase (early):** The agent reminds the team that it is early still, so exploration should be valued more.
5. **Notify phase (late):** The agent reminds the team that it is getting late, so exploration should be valued less.
6. **Remind of best practices:** The agent suggests that someone help a player who has a number of outstanding, but unaddressed, requests.
7. **Prompt activity:** The agent asks about any possible issues upon observing that the team has been predominantly idle for a period of time.

INTERVENTION-AC DEPENDENCY

Our ASI uses a dynamic influence diagram model that captures the dependency of the observable (via ACs) and unobservable teamwork variables. Our ASI uses this model (in combination with existing domain-independent algorithms) to make inferences about the unobservable team-process variables given input from the ACs. It then uses these inferences to inform its own decisions about team-level interventions by using existing domain-independent algorithms to compute the expected reward from each intervention. We can thus evaluate both the accuracy of the inferences (which may be difficult if ground truth is not available) and its interventions (which is already part of the evaluation plan, described in a later section).

The first source of interdependency, in this case of interventions on AC variables, is the trigger condition for each intervention. The following conditions must hold for the corresponding intervention to even be considered by the ASI, although they are not sufficient to actually trigger the intervention:

1. **Reflection:** Applicable only if the Rutgers AC reports a *wait_time* value over 15 seconds (the median value from early dry-run trials). If chosen, the agent reports on the incident with the maximum *wait_time* value.
2. **Cheerleading:** Applicable only if IHMC AC reports a joint activity whose *Activity* is “move-victim-to-triage-area” and whose *State* is “completion”.
3. **Report drop:** Applicable only if the Cornell AC reports *compliance_overall* for a requestor making at least one request of a requestee, who has not responded to any. The agent addresses this intervention to any player identified as an emergent leader (i.e., Gallup GELP AC reports a *Leadership* value higher than those for the other players) or to the team, if no emergent leadership is identified.
4. **Notify phase (early):** Applicable 5 minutes into the mission
5. **Notify phase (late):** Applicable 10 minutes into the mission
6. **Remind of best practices:** Applicable only if Cornell AC reports *compliance_overall* for a requestor waiting on at least five requests from teammates.
7. **Prompt activity:** Applicable only if CMU TED AC reports *inaction_stand_s* more than 5 seconds for more than 10 consecutive periods.

In addition to using the AC variables to determine the applicability of interventions, our ASI also models AC variables as evidence of the hidden team-process variables by drawing a probabilistic link between them in its influence diagram. The following variables are used to provide a conditional prior probability on the team-process variables as follows:

- CMU BEARD:
  - *anger*: negative influence on *affect management*
  - *anxiety*: negative influence on *affect management*
  - *gaming_experience*: positive influence on *systems monitoring*
  - *RMIE*: positive influence on *team monitoring*
  - *SBSOD*: positive influence on *systems monitoring*
- UCF:
  - *team-potential-category*: positive influence on team monitoring, coordination
  - *task-potential-category*: positive influence on systems monitoring

During the course of the interaction, our ASI uses the CMU TED AC to update its beliefs about the state of the team process variables. In particular, it includes the following links in its influence diagram:

- CMU TED:
  - *comms_equity:* positive influence on *coordination*
  - *comms_total_words*: positive influence on *team monitoring*
  - *inaction_stand_s*: negative influence on *systems monitoring*
  - *process_effort_agg*: positive influence on *motivating*

Given these links, our ASI can use standard Bayesian reasoning algorithms to update its beliefs about the team-process variables whenever it receives a new message from CMU TED. It does not use any other testbed messages, other than for reading the mission timer.

INTERVENTION-TEAM PROCESS INTERDEPENDENCY

The ASI’s goal (as reflected in its reward function) is to increase the team-process variables. We are starting with a uniformly weighted reward function that values all team-process variables equally, although this is unlikely to be the best formulation. Once data collection is complete we can refine the weights on this reward function to improve the intervention-selection outcomes. PsychSim provides some automated algorithms for potentially fitting these reward weights to a desired intervention strategy (Pynadath & Marsella, 2004).

To incentivize our ASI to choose its interventions, we introduce a positive effect on the team-process variables into our influence diagram. This positive effect is captured as a high probability of increasing the corresponding process variable. Our initial ASI models the interventions as improving team process variables as follows:

1. **Reflection:** *coordination*
2. **Cheerleading:** *motivating, affect management*
3. **Report drop:** *team monitoring*
4. **Notify phase (early):** *systems monitoring*
5. **Notify phase (late):** *systems monitoring*
6. **Remind of best practices:** *coordination*
7. **Prompt activity:** *team monitoring*, *coordination*

None of the interventions are modeled as having a negative impact on team-process variables, which is undoubtedly an oversimplification, but it is a useful first step. If no intervention is performed, then there is a low (but nonzero) probability that the process variables will increase. With only these reward components, the ASI would always want to intervene, so we introduce a disincentive in the form of a negative reward (an impact on *cognitive load*, in this case) for each intervention. We also introduce a flag that tracks whether the ASI has already chosen a particular intervention and forbid it from ever choosing it a second time in the same trial (although it can re-use the intervention for the same team in a subsequent trial). This is another obvious oversimplification, but we use it to eliminate the possibility of overly repeating interventions without having any data (or even a good guess) about how often such interventions would potentially be triggered.

Given all of these components, our ASI can compute the expected reward of any applicable intervention using existing domain-independent algorithms (Kaelbling et al., 1998). It is important to note that, although this initial model will most surely require significant revision, the algorithms operating it can remain unchanged. It is also important to note that, if any of the ACs were to drop out, our ASI would still be able to update its beliefs using whatever messages it did receive, and it would still choose interventions to maximize its expected reward subject to its beliefs.

EVALUATION

The program dictated high-level metrics for evaluation purposes. We present them and our specific implementation of them as follows:

1. AC Influence: Our ASI uses a maximum expected-utility calculation to choose its interventions. We can conduct counterfactual queries at each decision point as to whether a different input from a given AC would have prompted a different decision by our ASI. In fact, we will provide a quantitative measure of how much such a different input would have changed the expected utilities as a measure of the influence that the given AC's actual input had on the decision.4
2. AC Effectiveness: We can combine our quantitative AC Influence measures from #1 with weights based on the overall effectiveness of our ASI's interventions to provide a measure of effectiveness.
3. Interventions: We will count the chat messages generated by our ASI.
4. Compliance: Some of the ASI interventions are not relevant for compliance (e.g., cheerleading). However, compliance with its "distribute workload" intervention can be measured by changes in CMU TED's collective effort and skill use. We can measure it both in terms of aggregate increases,but perhaps more interestingly in equity of distribution (i.e., more balanced effort and skill use across players). Compliance with our ASI's "reminder about best practices" intervention can be measured by anticipated increase in CMU TED's communication.
5. Explanations: Our ASI's expected-utility calculations are in the form of explicit forecasting of expected effects of alternate courses of action. It can generate an explanation by comparing the differences across those alternate futures and providing the source of the biggest utility gains for the chosen intervention as a function of its current inferences about the team (e.g., the team was exhibiting poor coordination, which would be improved by reminding them of best practices).

6. MToT Existence Proof: Our ASI can provide a time series of its inferences over team-process variables: mission analysis, goal specification,strategy formulation, monitor goal progress, systems monitoring, team monitoring backup, coordination, conflict management, motivating, and affect management.
7. Inference: As mentioned, our ASI can function even if some ACs are not available. We can do a post-hoc evaluation where we re-run our ASI on modified logs where we drop the messages of a single AC. We can then compare the inferences made by our ASI about those AC variables against the values in the actual logs.
8. Prediction: As part of its expected-utility calculations, our ASI forecasts expected effects on the hidden team-process variables, but also the AC-provided variables. We will thus use those forecasts to generate predictions of the AC variables being used and compare them against the actual values eventually reported.

CONCLUSION AND FUTURE WORK

Data collection is currently underway at Arizona State University. The trials being run include not only our own ASI, but also the ASIs developed by other performers on the program, not to mention a control with no advisor at all, and a baseline with a human advisor. The challenge of coordinating joint participation of three participants has slowed the process down, so as of writing, five teams have run through the testbed with our agent. While there is obviously nothing conclusive that can be stated from this partial sample, a quick sanity check shows that the agent is choosing different interventions for different teams and even for different trials with the same team.

Once data collection is complete, we can perform the evaluation steps outlined above. Given that the current model was created through manual input (and with a great deal of guesswork, despite the combined expertise of the team), the more important contribution of these data will be in the refinement of the model. In particular, we can use the data to refine the dependencies captured in the current PsychSim influence diagram. Our prior work on fitting algorithms is one way to automate this process (Pynadath & Marsella, 2004). Inverse Reinforcement Learning (IRL) is another method we have applied in the past to extract reward functions from data (Sequeira & Gervasio, 2020). IRL is especially useful here in that it can provide clusters of reward functions as a way of creating broader categories of team members (e.g., people who value individual task achievement vs. people who value team process).

One early eye-opener from the data collected so far is that there was no significant improvement in team performance with a human advisor. As a result, we are tempering our expectations as to what our very limited ASI will be able to achieve and focusing on analyses that can inform the subsequent versions. Regardless, our current ASI represents a contribution in the form of a unified agent platform for incorporating team-process variables, social-science analyses from outside researchers, and (in the near future) human behavior data.

ACKNOWLEDGEMENTS

The authors thank the research teams at CMU, Cornell, Gallup, IHMC, Rutgers, and UCF who developed the ACs that serve as input to our agent. Thanks also go to Nancy Cooke and her team at Arizona State University for their ongoing work on collecting data for the evaluation of the agent. The experiment was approved by all of the performers' Institutional Review Boards (IRBs) and the US Army Human Research Protections Office (HRPO). This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. W911NF-20-1-0011. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Defense Advanced Research Projects Agency (DARPA).

REFERENCES

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2):241-251.

Bendell, R., Williams, J., Fiore, S.M., Jentsch, F. (2021). Towards Artificial Social Intelligence: Inherent Features, Individual Differences, Mental Models, and Theory of Mind. In: Ayaz, H., Asgher, U., Paletta, L. (eds) *Advances in Neuroergonomics and Cognitive Engineering. AHFE 2021. Lecture Notes in Networks and Systems*, Vol 259.

Corral, C.C., Tatapudi, K.S., Buchanan, V., Huang, L., Cooke, N.J. (2021). Building a Synthetic Task Environment to Support Artificial Social Intelligence Research. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 65(1):660-664.

Freeman, J., Huang, L., Wood, M., Cauffman, S. (2021). Evaluating Artificial Social Intelligence in an Urban Search and Rescue Task Environment. In *Proceedings of the AAAI Fall Symposium on Computational Theory of Mind for Human-Machine Teams*.

Gmytrasiewicz, P. J., & Doshi, P. (2005). A framework for sequential planning in multi-agent settings. *Journal of Artificial Intelligence Research*, 24:49-79.

Gurney, N., Marsella, S.C., Ustun, V., Pynadath, D.V.. (2021). Operationalizing Theories of Theory of Mind: A Survey. In *Proceedings of the AAAI Fall Symposium on Computational Theory of Mind for Human-Machine Teams*.

Hackman, J. R., & Wageman, R. (2005). A theory of team coaching. *Academy of management review*, 30(2):269-287.

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5):425-447.

Howard, R. A., & Matheson, J. E. (2005). Influence diagrams. *Decision Analysis*, 2(3):127-143.

Johnson, M., Hofmann, K., Hutton, T., & Bignell, D. (2016). The Malmo Platform for Artificial Intelligence Experimentation. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 4246-4247.

Johnson, M., Wang, N., & Pynadath, D. V. (2019). Simulating Team Tutoring in Multiagent Environments. In *Proceedings of the Approaches and Challenges in Team Tutoring Workshop at the International Conference for Artificial Intelligence in Education*, pp. 87-96.

Kaelbling, L.P., Littman, M.L., and Cassandra, A.R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1):99-134.

Maese, E., Diego-Rosell, P., DeBusk-Lane, L., Kress, N.. (2021). Development of Emergent Leadership Measurement: Implications for Human-Machine Teams. In *Proceedings of the AAAI Fall Symposium on Computational Theory of Mind for Human-Machine Teams*.

Marks, M. A., Mathieu, J. E., Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, 26(3):356–376.

Mathieu, J. E., Luciano, M. M., D'Innocenzo, L., Klock, E. A., & LePine, J. A. (2020). The development and construct validity of a team processes survey measure. *Organizational Research Methods*, 23(3):399-431.

Pynadath, D. V., & Marsella, S. C. (2004). Fitting and compilation of multiagent models through piecewise linear functions. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 1197-1204.

Pynadath, D. V., & Marsella, S. C. (2005). PsychSim: Modeling theory of mind with decision-theoretic agents. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1181-1186.

Pynadath, D. V., Si, M., & Marsella, S. C. (2011). Modeling theory of mind and cognitive appraisal with decision-theoretic agents. In *Social emotions in nature and artifact: emotions in human and human-computer interaction*, Gratch, J. and Marsella, S.C. (eds), 70-87.

Pynadath, D. V., & Tambe, M. (2002). The communicative multiagent team decision problem: Analyzing teamwork theories and models. *Journal of Artificial Intelligence Research*, 16:389-423.

Sequeira, P., & Gervasio, M. (2020). Interestingness elements for explainable reinforcement learning: Understanding agents' capabilities and limitations. *Artificial Intelligence*, 288:103367.