Proposal Submitted to the Air Force Office of Scientific Research
Trust and Influence program

Program Manager: Dr. Ben Knott

IMPROVING SITUATION AWARENESS IN DISTRIBUTED HUMAN-ROBOT TEAMS

October 31, 2016

Personnel

Nancy J. Cooke (PI)[1] & Subbarao Kambhampati (Co-PI)[2]

Nathan McNeese, PhD[1] (Co-I)
Yu Zhang, PhD[2] (Co-I)
Erin Chiou, PhD[1] (Co-I)

Human Systems Engineering[1] & Computer Science & Engineering[2]
Arizona State University

Mica Endsley, PhD, SA Technologies

Technical POC:

Nancy J. Cooke
Professor
Human Systems Engineering
Arizona State University
Phone: (480) 727-5158
Email: ncooke@asu.edu
https://webapp4.asu.edu/directory/person/559491

TABLE OF CONTENTS

1.0 Executive Summary

2.0 Problem Statement

3.0 Statement of Objectives and Timeline

4.0 Background

5.0 Related Work

6.0 Technical Approach: Collaboration

7.0 Technical Approach: Team Cognition

8.0 Technical Approach: Representations & Algorithms for Human-Robot Teams

9.0 Technical Approach: Human Systems Engineering

10.0 Technical Approach: Long-term Teaming and RPAs

11.0 Significance of Contribution

12.0 References

13.0 Personnel (bios)

14.0 Facilities

15.0 Special Test Equipment

16.0 Equipment

17.0 High Performance Computing Availability

18.0 Budget & Justification

PROJECT NARRATIVE

1.0 Executive Summary

There are many applications in which humans and robots need to team at a cognitive level to accomplish a task *too dull, dirty*, or *dangerous* for all-human teams. For example, robots can be used to speed rapid repair of bombed runways to maintain forward base operations or can serve as an unmanned wingman to a manned F-35. Often this cognitive teaming arrangement is not seamless and can compromise the human team member's workload and situational awareness, leading to slower decision-making and poorly calibrated trust between the human and the autonomy (USAF, 2015).

The interaction between human and robots differs in many ways from human interactions and these differences can contribute to poor teaming. Human teammates not only have teamwork skills (e.g., communication, back-up behaviors, conflict management), but they also are able to understand and relate to the actions of other teammates. They are able to recognize the intentions of team members, and to interact with them in a way that makes sense (i.e., is "explainable"). This basic ability to relate to human teammates may be especially important when there is minimal ability to communicate (e.g., communicating underground, denied and contested environments, limitations of natural language processing on the part of the robot).

We propose a transdisciplinary effort drawing from cognitive science (team cognition), artificial intelligence (robotic planning), and human systems engineering to improve situation awareness in such settings (see Figure 1). Deep knowledge integration across these disciplines will result in the most innovative science and solution to the problem. The extant literature on team cognition coupled with studies of human teams will inform robotic algorithms. A human systems engineering approach will evaluate the resulting robotic algorithms in terms of ability to team successfully with humans and preserve team situation awareness. Option years will be dedicated to extending this work over time (i.e., longer term human-robot teaming) and to the Remotely Piloted Aircraft System context.

The context for this research in the first three years is human-ground robot teaming of the type that might take place to repair damaged runways recently attacked or to do reconnaissance within a collapsed or dangerous structure. In each case one or more humans may work in concert with one or more robots in a distributed fashion. We use the game Minecraft to serve as a laboratory testbed for abstract scenarios that require the same kind of cognitive teaming that is needed in these real human-robot teaming scenarios. This context is selected for several reasons:

1) It exemplifies distributed human-robot cognitive teaming
2) We have several research testbeds suited to this context and have collected data in them
3) Because the scenarios we have used in the Minecraft task involve an uncertain environment they have the potential to tax situation awareness
4) The team interaction is less structured than some command-and-control tasks such as team control of remotely piloted aircraft
5) The USAF (2015) vision includes air and ground vehicles

This project will generate robot planning algorithms that are well-suited for teaming with humans in that they provide the essential behaviors necessary for effective cognitive teamwork and the preservation of team situation awareness. At the same time the science of team cognition will benefit from the identification of individual contributions to team effectiveness and team situation awareness especially under degraded communication conditions.

The proposed research is timely as it resonates with the increased interest in human-aware AI systems. The recent White House report (2016) on "*The National AI R&D Strategic Plan*" recommends developing effective methods for Human-AI collaboration as an important strategic direction, and recommends seeking new algorithms for human-aware AI. The proposed research is also in broad concordance with the AFOSR report on "*Recommendations for Research on Trust in Autonomy*" (Gratch, Friedland & Knott, 2016). In particular, we consider how to make the machines (robots) develop and maintain mental models of the human collaborators, and how to make them behave in such a way as to adapt to the human user's expectations, and, in long term, engender trust in the users.

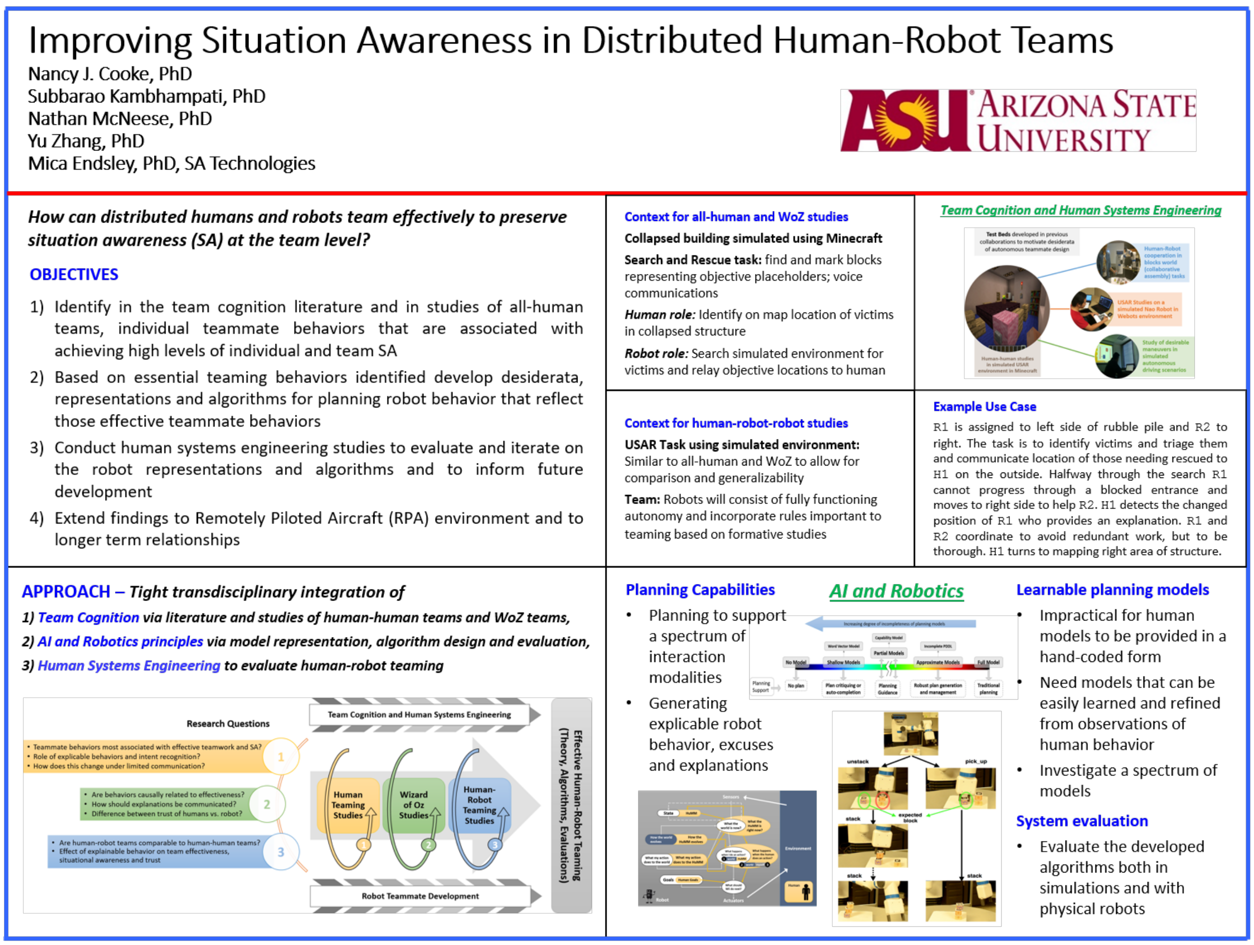


Figure 1. Quad chart summarizing proposed work

2.0 Problem Statement

What are the essential cognitive characteristics of robots required for team situation awareness and effective teaming in distributed human-robot teams? What representations and algorithms are needed to enable robots to realize such behaviors?

An increasing number of applications demand that humans and robots work together. Although a few of these applications can be handled through 'teleoperation', the USAF envisions technology with increasing levels of autonomy that act in concert with the humans in a teaming relationship (USAF, 2015). For instance, the USAF envisions conducting missions with humans teamed with decision aids and air and ground vehicles of varying levels of autonomy. In these types of situations, the teaming is primarily cognitive, rather than physical. Often this cognitive teaming arrangement is not seamless and can compromise the human team member's workload and situational awareness, leading to slower decision-making and poorly calibrated trust between the human and the autonomy (USAF, 2015).

The interaction between human and robots differs in many ways from human-human interactions and these differences can contribute to poor teaming. Human teammates not only have teamwork skills (e.g., communication, back-up behaviors, conflict management), but they also are able to understand and relate to the actions of other teammates. They are able to recognize the intentions of team members, and to interact with them in a way that makes sense (i.e., is "explainable"). This ability to relate to human teammates may be especially important when there is minimal ability to communicate (e.g., communicating underground, denied and contested environments, limitations of natural language processing on the part of the robot). We will conduct research to understand the essential cognitive characteristics of robots required for team situation awareness and effective teaming in distributed human-robot teams.

3.0 Statement of Objectives and Timeline

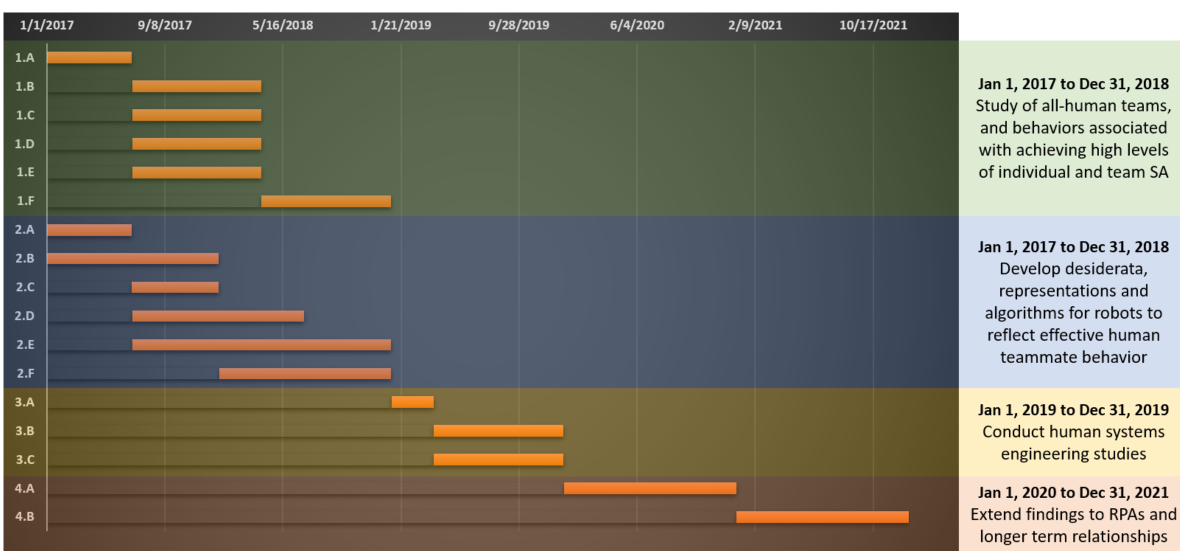


Figure 2. Gantt chart illustrating statement of work and timeline.

JAN 1, 2017 – DEC 31, 2018

1. **1.** Identify in all-human teams, individual Identify in the team cognition literature and in studies of all-human teams, individual teammate behaviors that are associated with achieving high levels of individual and team situational awareness (SA)
   - **A.** Develop scenarios and metrics (performance, SA, trust, workload) for teams using the Minecraft testbed [1/1/17-6/30/17]
   - **B.** Conduct a study in this context –with both all-human teams - to identify behaviors associated with high levels of individual and team SA [7/1/17-3/31/18]
   - **C.** Compare full communications to limited communications [7/1/17-3/31/18]
   - **D.** Attend especially to intent recognition and explainable behaviors [7/1/17-3/31/18]
   - **E.** Identify the essential vs. non-essential behaviors [7/1/17-3/31/18]
   - **F.** Test hypothesized behaviors derived in a Wizard-of-Oz paradigm [4/1/18-12/31/18]

JAN 1, 2017-DEC 31, 2018

2. Based on essential teaming behaviors identified develop desiderata, representations and algorithms for planning robot behavior that reflect those effective teammate behaviors.
   - **A.** Identify the challenges involved in synthesizing robot behavior in the presence of humans [1/1/17-6/30/17]
   - **B.** Develop algorithms for tracking the intent of the humans that can work with shallow planning models [1/1/17-12/31/17]
   - **C.** Develop algorithms for planning with humans to support collaboration [6/30/17-12/31/17]
   - **D.** Develop algorithms for generating explicable behavior (which involves tracking human's models of the robot's capabilities) [7/1/17-6/30/18]
   - **E.** Develop mechanisms where by the robot can signal its intent when explicable behaviors are too costly [7/1/17-12/31/18]
   - **F.** Generalize the robot teaming algorithms to account for multiple robots [1/1/18-12/31/18]

JAN 1, 2019-DEC 31, 2019

3. Conduct human systems engineering studies to evaluate and iterate on the robot representations and algorithms and to inform future development
   - **A.** Integrate the implemented team behaviors [1/1/19-3/31/19]
   - **B.** Test in scenario with one human and two cognitive robots in search and triage task [4/1/19-12/31/19]
   - **C.** Evaluate and iterate [4/1/19-12/31/19]

OPTION: JAN 1, 2020-DEC 31, 2021

4. Extend findings to Remotely Piloted Aircraft System environment and to longer term relationships

A. Study of one human and two RPAs or two humans interacting with synthetic teammate to control one RPA [1/1/20-12/31/20]

B. In search and triage domain bring humans back to work with same robots and examine dynamics of trust over time [1/1/21-12/31/21]

4.0 Background

In this section, we review our own work and that of others to position this project in the team cognition and robotics literatures.

4.1 Teams and Team Cognition

Many complex tasks require teams of interdependent individuals with different skill sets to work together to achieve a shared goal. The diverse roles and interdependence are the essence of what the term "team" means (Salas, Dickinson, Converse, & Tannenbaum, 1992). Teams can conduct physical and cognitive tasks together and in this project, we focus on the cognitive tasks such as planning, sensing, communicating, deciding, and problem solving. In many settings, these tasks require tight coordination and interdependent teamwork.

There have been two primary perspectives on team cognition that can be seen as complementary. The shared cognition or shared mental model perspective frames team cognition as an aggregate of the knowledge or mental models of individual team members (Cannon-Bowers & Salas, 2001). Teams whose members share or have similar mental models are thought to have better implicit coordination and more effective teamwork (Entin & Serfaty, 1999).

The shared mental model perspective relies on measurement of knowledge at the individual level and aggregation to estimate team-level knowledge. It also provides a relatively static snap shot of team cognition. Another perspective which focuses measurement at the team level and provides a more dynamic view is interactive team cognition (Cooke, Gorman, Myers, & Duran, 2013). This perspective views team cognition as the cognitive processing that is done by teams through various forms of interaction – often in the form of verbal communication. In various settings with action-oriented teams (e.g., military command-and-control) interactions and interaction dynamics have accounted for team effectiveness.

In other settings where teams do tasks that rely more on shared knowledge such as design teams or science teams, then team member knowledge may better account for performance. In general, it is valuable to keep both perspectives in mind when determining how to measure team cognition or how to intervene to improve it. Situation awareness, for instance, according to the shared mental model view can be conceptualized as a team model (or knowledge) of the situation based on the aggregate of individual team member models. On the other hand, from the interactive team cognition perspective, team situation awareness would be conceptualized as the interaction that occurs as teams assess a situation and take action (Gorman, Cooke, & Winner, 2006). In the next two sections we describe how teaming concepts from human studies can be applied to human-robot interactions and then how SA specifically might be considered.

4.2 Human-Robot Teaming

Human-robot teaming is becoming increasingly important in a variety of tasks, such as search and rescue, and command and control (Kruijff et al., 2012; Parasuraman, Barnes, Cosenzo, &

Mulgund, 2007). There are multiple instances of human-robot teams successfully teaming together to complete their objectives (e.g. Casper & Murphy, 2003). Yet, there are also currently many limitations associated with human-robot teaming that can be viewed as barriers to effective teaming, specifically cognitive teaming. The nature of human-robot teaming can lead to limited communication, increased human team member workload, and a lack of situational awareness, all of which limit team effectiveness (USAF, 2015). In response to human robot-teaming limitations, research is needed to fully understand how these teams work together effectively.

A great deal of knowledge is known about human-human teaming. An in-depth understanding of how human-human teams plan, communicate, and make decisions has been derived over the past 30 years (Salas, Cooke, & Rosen, 2008). Through these insights, we know the characteristics (e.g., role clarity, situational awareness, shared knowledge, closed loop communication) that lead to effective human teaming. In the human-robot teaming context we need to leverage our knowledge of effective human-human teaming to verify and translate how human-human team level characteristics may occur in human-robot teams (Wiltshire, Barber, & Fiore, 2016). Observations of distributed human-human teamwork, though not prescriptive for human-robot teaming, can provide hypotheses for improving human-robot teaming.

Like many human-human teams, human-robot teams are often distributed. That is, the human team member and the robot team member are separated by space and in some cases time. Distributed teamwork has been studied in detail in human-human teaming, and the findings from human-human teams may provide insights into bettering human-robot teaming. For example, many of the factors affecting technology-mediated interactions between people; such as trust, communication, and work organization factors; are not exclusive to human-human teams (Montague & Chiou, 2014). Furthermore, as robot capability increases they may begin to take on roles similar to distributed human teammates, requiring a better understanding of potential social influences in coordination like intent recognition (Chiou & Lee, 2016).

In general, distribution of team members, whether human or not, can lead to a multitude of issues that negatively affect teaming behaviors and performance. Issues specific to temporality, geographic dispersion, and varying team configurations may lead to a lack of team level communication and coordination (O'Leary & Cummings, 2002). In addition, there are issues with the use of technology in distributed team settings. We know that collaborative technologies may create a barrier between both team members and their environment. More specifically, the concept of team opacity, the loss of team situational awareness through technologically mediated communication, can lead to a lack of team situation awareness (Fiore, Salas, Cuevas, & Salas, 2003). Similarity, Stagl and colleagues (2007) have outlined issues pertaining to the use of collaborative technologies limiting situation awareness in distributed teams. In addition, research has also indicated that shared knowledge is limited in distributed human-human team settings. Work conducted by Bolstad and Endsley (1999) identified that shared mental models developed faster in settings where team members were physically close as compared to team members who interacted through a distributed technology. Similar, McNeese and Reddy (2015) identified that shared mental models developed in a slower manner and were confined to only taskwork (and not teamwork) characteristics due to multiple communication issues. Cooke and colleagues (Cooke, Gorman, Myers & Duran, 2013) have outlined in their theory of interactive team cognition the importance of communication to teamwork. This work, taken in concert with the work outlining team communication issues in human-human distributed teams helps to explain issues of team effectiveness in distributed settings.

Communication is often minimal in human-robot teams due to limits of natural language processing, denied or contested environments, or limited communication bandwidth in extreme environments. The communication deficits that occur in human-robot teams will not only impact the logistical components of teaming, but will also affect variables such as situation awareness and short and long term trust especially when team members are not human (for more on trust and in human agent interaction see Gratch, Friedland, & Knott, 2016). Communication is clearly fundamental and often a hindrance to effective human-human distributed teaming, but there are other types of distributed teams that are limited in their communication and still perform at high levels. Recently, researchers have highlighted human-dog teams as an analog for understanding human-robot teams (Kapalo, Phillips, & Fiore, 2016; Gratch, Friedland, & Knott, 2016). Human-dog teams are often used in a variety of contexts and have shown to be effective with limited communication abilities. More research into the effective and ineffective communication instances in human-dog teams is necessary to be translated fully to human-robot teams.

Although the teamwork literature has identified characteristics of effective and ineffective teams and their members (e.g., role clarity, closed-loop communications) it is less clear what characteristics are essential at the individual level to be a good team player. What are the most important behaviors that a teammate must exhibit for the team to be effective? For instance, is closed-loop-communication essential or is it more important to develop a good understanding of team members' intentions and capabilities? What exactly does a robot need to understand about its human teammates? The relative importance of these behaviors is not easily isolated in experiments with human teammates, but can be isolated in robot models. Neither is it understood as to what representations and planning algorithms are needed to support such behavior. Our research team has expertise in human teaming and in robot models and therefore is well-positioned to conduct research to identify the critical features associated with being an effective teammate. More specifically, we intend to focus on individual and team situational awareness (SA) in human-robot teams operating with limited communication bandwidth. SA is noted for its importance in human-human teaming and we need to better understand how it occurs in human-robot teaming to improve human-robot team effectiveness.

4.3 Situation Awareness and Human Robot Teaming

In general, team effectiveness is tied to shared knowledge (Mathieu et al., 2000), interaction (Cooke, Gorman, Myers, & Duran, 2013), and situational awareness (SA) (Gorman, Cooke, Winner, 2006). Therefore, we assume that robots that are effective team players will need to have some capabilities for understanding their teammates and the situation and for communicating their actions to teammates. Specifically, individual and team SA are critical to teamwork and relates to all aspects of human-robot teaming behavior, from communicating effectively to sharing knowledge.

SA has a long history within the human factors community, being applied to multiple different contexts (e.g. aviation, command and control, healthcare; see Endsley, 2015 for a review of SA). Many interpretations of SA have been postulated over the years. In general, SA is defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley 1988, p. 97). This definition has been interpreted and applied as both a product and a process (Smith & Hancock, 1995). In theory, one could end up accumulating a state of knowledge representative of the elements in their environment (a product). Although, the position that SA is only a knowledge state (or product) might not be best for environments that are highly dynamic.

In a dynamic environment, elements (information) of the environment are constantly changing. These ever-changing environmental elements would imply that SA is a never finished product. In response to SA in dynamic environments, Gorman and colleagues (2006) have suggested that SA or situation assessment is a continuous perception-action process that relies on humans adapting to a perceived element of the environment. In general, SA is both a process and product, associated with adaptation to change in the environment.

Of particular importance to human-robot teaming is the concept of team situation awareness (TSA). Similar to team cognition, there are two perspectives of TSA that are complimentary to each other- a shared knowledge approach and an interactive approach. Traditionally, TSA is viewed from a shared knowledge perspective and defined as 'the degree to which every team member possesses the SA required for his/her job' (Endsley and Jones, 1997). Essentially, this requires the aggregation of each team member's individual SA to then represent TSA. This viewpoint suggests that individual SA for each team member is important, as well as the degree of shared SA (or shared knowledge on information needs that overlap across multiple team members). The shared overlap of these knowledge states then becomes SSA (shared SA), similar to the concept of shared mental models (Mohammed, Ferzandi, & Hamilton, 2010). This approach to TSA is highly valuable depending on the team's task and environment. If a team task is knowledge intensive then a shared knowledge approach to TSA is effective.

USAF (2015) identified SSA within human-autonomy teams as being critical for supporting key aspects of team performance including: dynamic goal alignment as priorities shift; function allocation and reallocation across team members based on changing capabilities and status; communication of decisions, strategies, plans and actions; and coordination of status and success on inter-related tasks. Coordination and collaboration between humans and robots to align world views in terms of SSA on these issues is required to achieve successful team performance.

Endsley and Jones (1997) emphasize four main contributors to TSA: a need for shared knowledge on shared SA requirements (SSA); shared mental models (which can make gaining shared SA of comprehension and projection easier); shared SA devices (which transmits need information between teammates); and effective team processes (e.g. communication and coordination). And more emphasis on one factor (shared mental models, shared devices or team processes) can be used to compensate for the loss of others in the development of TSA (Bolstad and Endsley, 1999).

Gorman and colleagues (2006) have suggested that TSA is directly linked to the interactions that occur as teams assess a situation and take action. This viewpoint of TSA is directed towards environments that require interaction in response to dynamic environmental elements. From this perspective, TSA requires communication and coordination of information to the right person at the right time (dependent on responses to the environment), consistent with the team processes aspect of the Endsley and Jones (1997) model of TSA. In addition, from this perspective, situation awareness at the team level does not equal the accumulation of each individual team member's SA.

We view both the shared knowledge approach and the interactive approach to TSA as being equally important depending on a team's tasks and environment. Also, it is possible for TSA to occur simultaneously in tasks/environments that are highly dependent on both knowledge and interaction. Most tasks and environments in human-robot teaming are highly dynamic with changing information, requiring team level adaptiveness. At the same time, human-robot interaction is constrained because they are often geographically separated with limited communication bandwidth. Under these circumstances, team SA may have to rely on both knowledge and interaction.

Additionally, robot behaviors may lack some very basic human capabilities. The robot algorithms may generate behaviors that appear odd to a human. When communication is limited, this unpredictable behavior may further the divide between human and robot teammates, impacting team SA. Therefore to increase the effectiveness of human-robot teams we must generate robots that take actions that are explainable and predictable. Specifically, these actions need to account for how they influence both individual and team SA. This raises the fundamental question, if TSA is highly dependent on communication, how do effective teams achieve high levels of TSA given limitations in the human-robot teaming context? Can shared mental models that provide predictability of actions in a given environment compensate for such limitations? This proposal seeks to provide answers to this question and apply those answers to generate explainable robot actions that are directed at enhancing both individual and team SA in this unique environment.

Previous research focusing on both individual and team SA in human-robot teaming can help provide a foundation for understanding how to achieve SA in human-robot teaming. For example, Riley and Endsley (2004) observed USAR professionals operating a single robot in a physical environment for a training mission. Through their observations, multiple SA challenges related to human robot interaction (HRI) (teleoperation/supervisory control) were identified including disorientation of the robot, interface issues not supporting SA, and the need for interfaces to specifically support TSA. In a separate study, Riley and Endsley (2005) explored collaborative HRI in which a single unmanned aerial vehicle (UAV) was coordinating multiple unmanned ground vehicles (UGV). Goal-directed task analyses and function allocation for the UAV and UGV tasks were identified, suggesting that human operators must develop and maintain SA in relation to 1) individual robots, 2) a group of robots, and 3) the status of the overall task. Connors and colleagues (Connors, Strater, Riley, & Endsley, 2008) conducted a goal directed task analysis on the collaborative control of multiple unmanned vehicles and generated requirements for collaborative technology to enhance SA. Their suggestions indicate that collaborative technologies need to promote shared SA by accounting for Level 2 (comprehension) and 3 (projection) of SA, the development of a common operational picture, and the need for critical information to be shared across technologies. Finally, Riley and colleagues (Riley, Strater, Sethumadhavan, Davis, Tharanathan, & Kokini, 2008) looked at collaborative robot control with semi-autonomy and found that being aware of robot direction and when a robot had completed a task were critical to SA. Findings from this study also suggest that robot interface designs must be complex and account for multiple informational details in order for the human to be situationally aware.

It is clear through the existing literature that achieving both individual and team SA in the HRI context is complicated. There are multiple issues relating to knowledge, communication, coordination, and technology that can impact both individual and team SA in a human-robot team. Although, the research presented here is very valuable in understanding individual and team SA in HRI, more work is needed to understand specific teamwork behaviors associated with high levels of individual and team SA during human-robot teaming in which communication is limited.

4.4 State of the Art – Human-Robot Teaming in AI and Robotics

Although there is a growing body of work on robots teaming with humans, very little of it focuses on the cognitive teaming scenarios of the kind we are interested. Specifically, much of the existing work on human-robot teaming focuses on *proximal teaming* mostly involving robot motion planning, such as robot navigating in crowded environment with humans (Kruse et al., 2013) or robot handing tools to human teammates (Shah et al., 2011; Unhelkar et al., 2014}. While important, planning for motions only represents a small part of human-robot teaming applications. More often, it is desirable for human-robot teams to perform complex tasks requiring cognitive teaming such as urban search and rescue, elderly assistance and etc. For robots to be helpful, they are required to automatically generate task plans based on high level goal specifications. In this research, when we refer to ``planning'', we refer to task planning. In the following, we discuss the challenges in planning for human-robot teams, and the few research efforts aimed at some of those challenges.

Human-aware Planning

Most traditional approaches to planning focus on one-shot planning in closed worlds given complete domain models. Although even this problem is quite challenging, and significant strides have been made in taming its combinatorics, planners controlling the co-robots in human-robot teaming scenarios require the ability to be *human-aware*. Human-aware planning is aimed at producing plans that consider the human models, which include, for example, people's capabilities and preferences. These models can be used to infer human intent and plans which in turn informs the robot's decision making. In the robotics and automated planning communities, there has been a lot of work recently in human-aware planning, both from the point of view of path planning in and task planning in (Talamadupula *et. al.,* 2010; Hayes & Scassallati, 2015, 2016) with the intention of making the robot's plans socially acceptable, e.g. resolving conflicts with the plans of human counterparts. However, robots in most of these approaches focus on maintaining a belief about the person's intent and plan and how to use this information for plan generation. For a more comprehensive survey of the existing methods for human-aware and human-in-the-loop planning, the reader is referred to our AAAI 2014 tutorial on the subject (Kambhampati & Talamadupula, 2014).

Explainable Plan Generation

Explainable plan generation represents the ``implicit'' part of communication in intent projection which allows robots to generate plans that are explicable and comprehensible to people. In the human-robot interaction community, there exists prior work that discusses how to enable natural and fluent human-robot interaction (Shah *et. al.* 2011; Unhelkar *et. al.* 2014) to create more socially acceptable robots (Fong *et. al.,* 2003). Griogore *et. al.* (2016) and Admoni *et. al.* (2016) demonstrate how a robot's verbal and non-verbal behavior can improve task performance. These discussions, however, apply only to specific behaviors and application domains, and thus cannot be generalized. In addition we see explainable plans as critical to the situation awareness of the human and consequently the formation of team situation awareness. The plan explainability and predictability measures we propose can be used by agents to proactively choose, or directly incorporate into the planning process to generate plans that are more explainable and predictable contributing to overall plan quality.

There exists work on generating legible robot motions (Dragan *et. al.* 2013) which considers a similar issue in motion planning. We are, on the other hand, concerned with task planning. Note that two different task plans may map to exactly the same motions which can be interpreted vastly differently by humans. In such cases, considering only motion becomes insufficient. Nevertheless, there exists similarities between Dragan *et. al.'s* work and our work on explainable plan generation. For example, legibility there is analogous to predictability in ours.

Incomplete and Learnable Planning Models

Shared mental models are a critical part of human-human teaming. We thus expect that effective human-robot teaming also requires the robots to track the mental models of the humans in the team. These models are inherently partial from the robot's perspective and must be learnable. However, most work on planning has hitherto focused on complete models (Ghallab et al., 2004; Geffner and Bonet, 2013; Kambhampati, 1997). Most real-world scenarios have a critical defining property that eliminates these standard planning and problem-solving methods: they are *open-ended* in that planning agents typically do not have sufficient knowledge about all task-relevant information at planning time -- in other words, the planning models would be incomplete. This is especially true in human-robot teaming because the robotic agents may only have partial information about the humans in terms of their capabilities and preferences. Hence, an important challenge is to develop representations of approximate and incomplete models that are easy to learn, and can support planning/decision-making.

In our earlier work, we started with complete action models, such as those in STRIPS representation, and modified them with possible precondition/effect annotations to support incompleteness (Nguyen & Kambhampati, 2014; Nguyen & Kambhampati, 2013; Zhuo, Nguyen & Kambhampati, 2013; Zhuo, Nguyen & Kambhampati, 2013b). Although these models support principled approaches for robust planning, they are still quite difficult to learn. We have also proposed shallower and easier to learn representations including *capability models* (Zhang, Sridharan & Kambhampati, 2015) and *action vector models* (Tian, Zhuo & Kambhampati, 2016) that assume no structured information at all which are used mainly in short-term planning support. In the proposed work, we plan to investigate the applicability of this spectrum models.

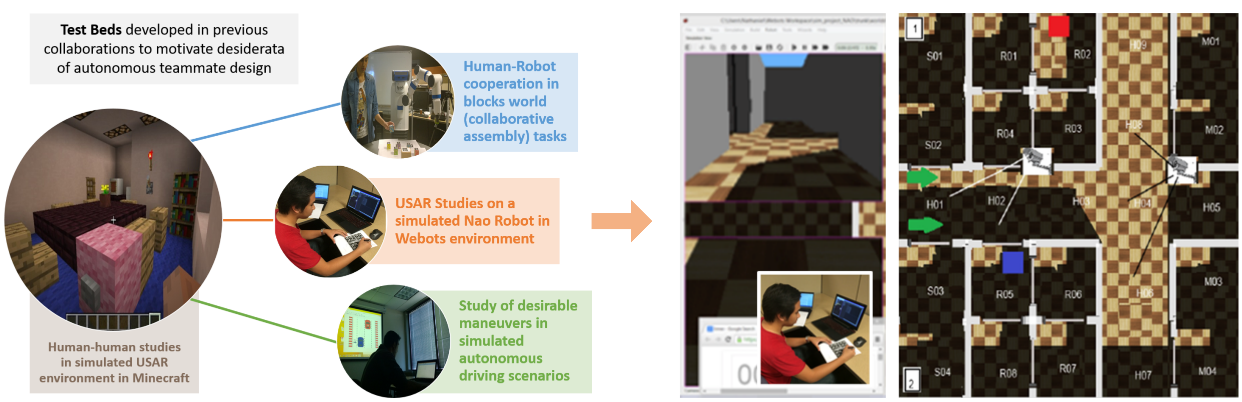


Figure 3. Testbeds developed in previous collaborations. The platform for simulated human-robot USAR studies in Webots is shown on the right.

5.0 Prior Collaboration between the PIs

Our previous successful collaboration of more than three years has led to a number of accomplishments that we are able to leverage for this work. They include:

- A Minecraft-based testbed for conducting human teaming and Wizard-of-Oz studies in a distributed cognitive teaming context (Bartlett & Cooke, 2015).
- A similar testbed for conducting human-robot interaction studies with Webots simulator (Narayanan, Zhang, Mendoza & Kambhampati, 2015; Zhang, Narayanan, Chakraborti & Kambhampati, 2015), as well as physical robotic platforms (Zhang et al., 2015) (see Figure 3)
- Three studies in a distributed cognitive teaming environment that lay some groundwork for this effort (Bartlett & Cooke, 2015; Narayanan, Zhang, Mendoza, & Kambhampati, 2015; Zhang, Narayanan, Chakraborty & Kambhampati, 2015).
- Analysis of distributed teaming data with human teams indicating that effective teammates are proactive in providing suggestions, often in the form of excuses, than ineffective teammates (Bartlett, McNeese, Cooke, Zhang, & Kambhampati, under revision).
- Ongoing collaboration of Kambhampati and Cooke labs on planning challenges inherent in human-robot teaming.

In addition, our research team brings years of scientific accomplishments that can be leveraged in this collaboration

- Twenty years of studies of teamwork resulting in a theory of team cognition (Cooke, Gorman, Myers, & Duran, 2013) and providing a starting point for hypotheses about effective human-robot teaming.
- Twenty five years of experience in automated planning methods, with recent thrusts in decision making with incomplete models for human-robot teaming, with studies in virtual as well as embodied teaming (Kambhampati's lab has several state of the art robots, including a Baxter, a Fetch, a PeopleBot and several NAO robots).
- Twenty years of studies on team SA and its measurement in human-human teams providing a foundation for studying human-robot teams, as well as 15 years of studies on human-robot teaming.
- Five years of experience studying trust in automation and three years of experience conducting related microworld studies to explore dynamic coordination and cooperation in human-agent teams with limited communication (Chiou & Lee, 2016; Montague & Chiou, 2014; Montague, Asan, & Chiou, 2013).

6.0 Technical Approach: Collaboration

We propose to conduct parallel threads of research on team cognition and modeling and algorithm research to support human-robot teaming. These parallel threads will be integrated through a series of studies in a distributed cognitive teaming task that progress from human-human teaming, to Wizard of Oz (WoZ) scenarios in which a person plays the role of the robot, to human systems engineering studies of human and autonomous robot teaming (see Figure 4). The team cognition thread will build an understanding of effective robot teammate behavior by conducting the experiments and iterating on teammate characteristics. The robot thread will develop robot representations and algorithms as informed by the empirical results of the studies. In addition, the human systems engineering studies will serve as evaluations of the robot teammate theories and the representations and algorithms.

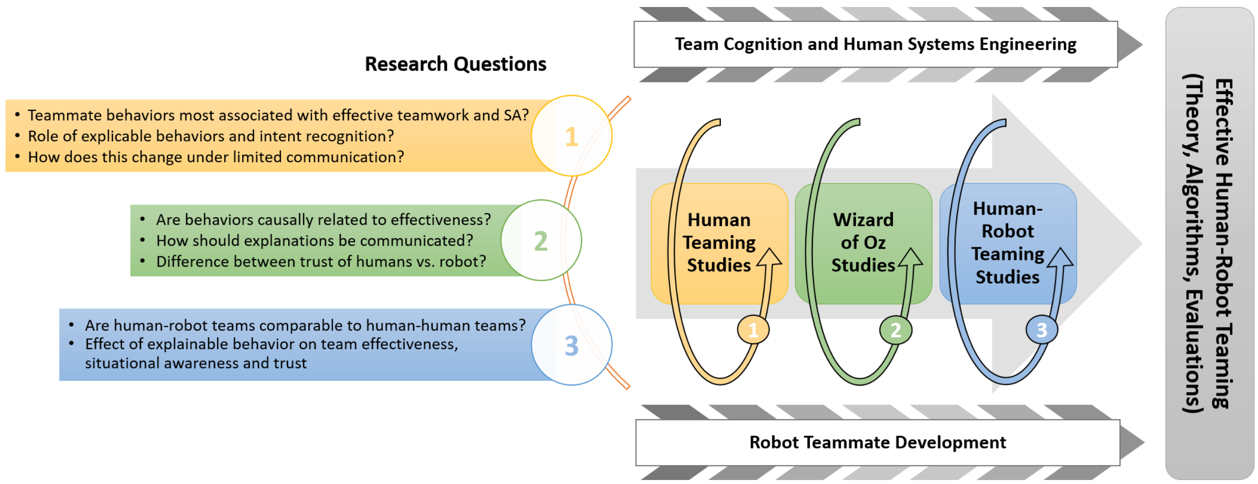


Figure 4. Tightly integrated collaboration between team cognition and robot planning research and human systems engineering.

7.0 Technical Approach: Team Cognition

The objective to be addressed though the literature and studies on team cognition is:

Identify in the team cognition literature and in studies of all-human teams, individual teammate behaviors that are associated with achieving high levels of individual and team SA.

The tasks that are necessary to achieve this objective are:

a. Develop scenarios and metrics (performance, SA, trust, workload) for teams using the Minecraft testbed
b. Conduct a study in this context –with all-human teams - to identify behaviors associated with high levels of individual and team SA
c. Compare full communications to limited communications
d. Attend especially to intent recognition and explainable behaviors
e. Identify the essential vs. non-essential behaviors
f. Test hypothesized behaviors derived in a Wizard-of-Oz paradigm

An initial literature review will be used to identify an initial set of teammate behaviors that are associated with team effectiveness. Metrics that are developed will target these behaviors.

Human Teaming Study

We begin with human-human teaming studies to understand how human teams work in the search and triage context and to identify individual factors leading to effective teamwork and situation awareness. We will also manipulate the degree to which verbal communication is allowed, simulating human-robot conditions. Research questions to be addressed in the first study are:

- What individual teammate behaviors are most associated with effective teamwork and team SA?
- What is the role and nature of explicable behaviors and intent recognition in effective teamwork and SA?

- How do these effective behaviors change under conditions of limited communication and what is the effect on trust and workload?

Scenario and Metric Development. The studies presented here will take place within the context of triage distributed cognitive teaming task that involves robot reconnaissance of a structure remotely located from the human teammate. This is a context that is highly reliant on robots and human-robot teaming (Kruijff et al., 2014). We will use the Minecraft platform (https://minecraft.net/en/) to simulate this context. This is a customizable platform that aligns well with the tasks and goals of human-robot teaming in tasks such as runway repair, and has successfully been used to conduct human-robot teaming research in the past (see Figure 5; Bartlett & Cooke, 2015).

Figure 5. Human-human study simulator.

At a high level, the task will consist of teams working together to find survivors and threats within a collapsed structure that was recently attacked. Team members are either internal or external to the actual environment. An internal team member does the work of the robot, whereas the external team members represent the humans out of harm's way on the outside of the structure. The external team member scopes the environment marking survivors and threats on a map (Figure 6), based on communication with the internal team member. The nature of this task requires team cognition and situation awareness to be successful. Effective communication and coordination will be necessary for the external and internal team members to find survivors and identify threats.

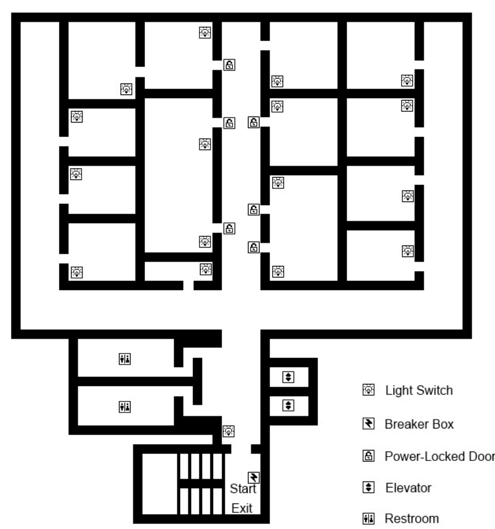


Figure 6. Map for external team

Using Minecraft and previous scenarios as a starting point, a scenario and Minecraft structure will be created and iteratively evaluated in pilot testing. The goal is to assure that the scenario has adequate levels of complexity and features capable of challenging SA and eliciting changes in plans and therefore the need for explanation. Complexity can be increased by having an additional internal team member (team of 3 vs. 2), by increasing spatial complexity (multiple floors, larger discrepancy between structure and map), by adding temporal complexity (e.g., need to identify certain survivors and threats by deadline), by increasing sources of information, or by adding a secondary task to the tasks of the external.

At the same time metrics will be adapted from previous studies and also tested in the new scenario context. Metrics will include task performance, individual and team situation awareness, workload and trust. A metric will also be developed that captures specific teammate behaviors that have been identified in the literature and by the robot planning team (intent recognition, explicable behaviors). Pilot testing will be conducted on scenarios and metrics which will be revised accordingly and tested again in an iterative fashion until objectives are met.

Experimental Method. Using the scenario and metrics developed we will conduct a study with 30 teams of 2-3 humans in which one team member is randomly assigned to be the external team member and the others are assigned to be internal. Half of the teams will be assigned to the full communications condition and the other half to limited communications in which they are allotted fixed vocabulary and two word utterances at a rate of no more than one utterance a minute. Internal team members will be seated at the Minecraft workstation and will be visually separated from each other and the external by a partition. The team members will be provided instruction sheets pertinent to their tasks and with a map of the structure pre-collapse. Participants will have five minutes to come up with a search plan and the internal participant(s) will be given brief instruction in Minecraft. Once this is complete the task will begin. The external participant will identify the location of survivors and threats on the map and the internal participant(s) will search for survivors and threats in the simulated environment and relay their locations to the external participant. Participants will have 15 minutes to complete the task. After the task is complete they will be administered the NASA TLX workload assessment and a questionnaire that assesses teammate trust.

Wizard of Oz Study

The Wizard of Oz paradigm involves a human playing the internal role of the robot, while the external human is told that they are interacting with a robot. The paradigm is particularly helpful when comparing different forms of "robot" behavior without having to develop different robot interfaces. In this study, we take the findings from the first study and test them by generating rules for "autonomy" to interact with humans. If certain teammate behaviors are found to be associated with effective teaming, we can directly manipulate them in this study to assess whether the relationship with effectiveness is a causal one. We can also compare two or more forms of interaction. For instance, explainable behavior may come with additional verbal input or not. This study will use the same scenario and metrics developed for the human teaming study. Similar to Study 1, we will measure team effectiveness, team situation awareness workload and trust. We will also measure communication and coordination in real time. One of the main goals of this study will be to test the effectiveness of the generated rules. In general, we will be seeking to understand if the rules (vs. no rules) produce more effective teams. We will then take these results and provide input into the development of autonomy for real robots. Specific research questions include:

- Are identified teammate behaviors causally related to team effectiveness and SA?
- How should explanations or excuses be best communicated for effective teaming?
- How do human interactions with and trust of robots differ from those with humans?

Through the human teaming and Wizard of Oz studies we will identify the essential requirements of a robot teammate for effective interaction, situation awareness and trust. This work will inform the robot development as well as the team cognition literature. Moreover, natural language communication would also help in making robot behavior more explainable and thus facilitate fluent human-robot teaming. Through our research we can identify the minimum language capabilities needed by a robot in this environment for effective teaming with humans. This information will inform language development efforts.

8.0 Technical Approach: Representations & Algorithms for Human-Robot Teams

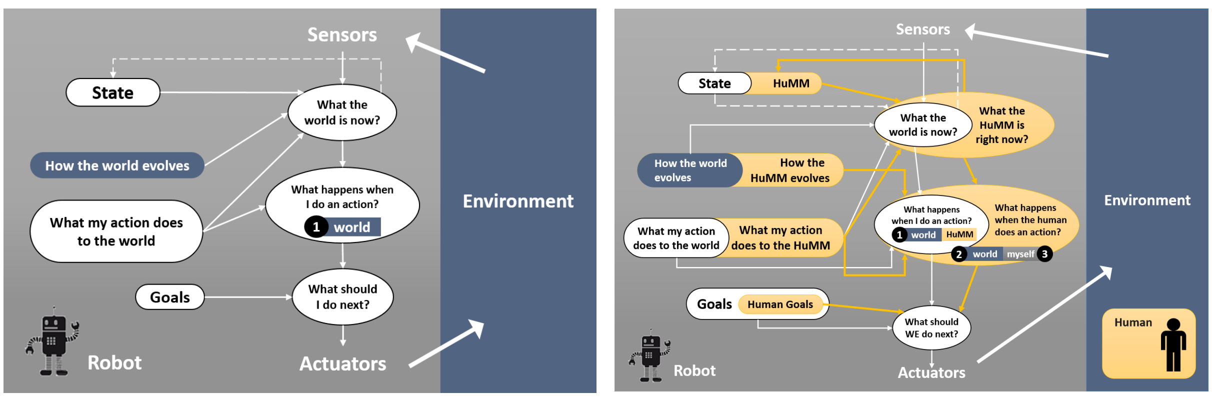


Figure 7. Human Mental Models (HuMM) are essential for natural human-robot teaming.

The artificial intelligence community has traditionally considered the external environment as a whole and attempted to model autonomous behavior based on a well-known model of an intelligent agent shown in Figure 6 (left). However, we argue that the presence of a human in the loop introduces significant and unique challenges in the deliberative process of the autonomous agent – it must now be able to model the beliefs, desires and intentions of its teammate and adapt its plan accordingly. We refer to this as the Human Mental Model (HuMM) and propose an alternative approach to the design of autonomy in Figure 6 (right).

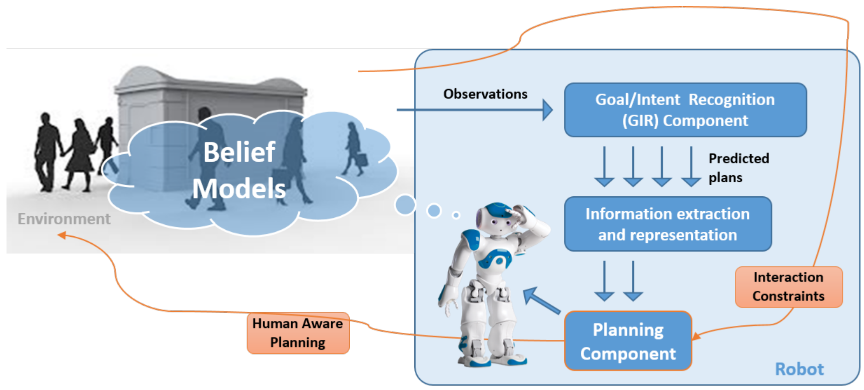
As we mentioned at the outset, our interest is to understand the cognitive characteristics of robots required for team situation awareness, and leverage that understanding in designing representations and algorithms for guiding the behavior of robots involved in human-robot cognitive teaming scenarios. This latter involves tackling several significant research challenges:

a. Identify the challenges involved in synthesizing robot behavior in the presence of humans
b. Develop algorithms for tracking the intent of the humans that can work with shallow planning models
c. Develop algorithms for planning with humans that support collaboration
d. Develop algorithms for generating explicable behavior (which involves tracking human's models of the robot's capabilities)

e. Develop mechanisms where by the robot can signal its intent when explicable behaviors are too costly
f. Generalize the robot teaming algorithms to account for multiple robots

In the following we briefly discuss the technical approaches we intend to pursue for tackling these challenges.

Human-Aware Planning for Human-Robot Cohabitation



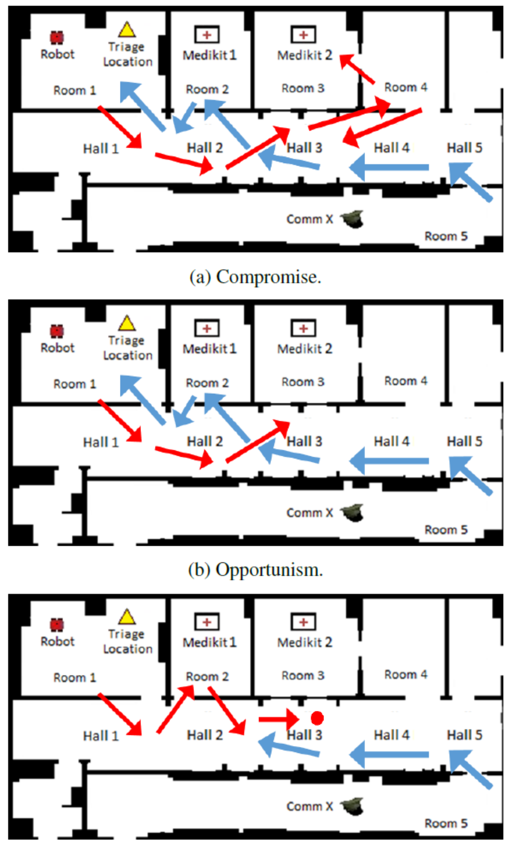
When robots operate with human teammates, it is often not enough to just plan for achieving individual sub-goals optimally. They must, for example, determine if their assistance would be considered useful from the human's perspective at all, or whether its plans are interfering with that of the human's expected plan in terms of resources used. This aspect of planning, referred to as human-aware decision making, is becoming increasingly important as human-robot teams proliferate. To support such behavior, the robot needs to model the humans in terms of their capabilities, intentions and preferences, and also remain robust to the incompleteness of the human models, in terms of the uncertainty in human goals and intent. Hence, human-aware decision making requires developing algorithms that can make interaction among humans and robots in collaborative tasks closer to that expected naturally among human teammates.

While it is well understood that planning technologies for robots must be "aware" of the humans sharing their workspace, different types of human-robot teams (in terms of proximal vs remote teams, to the amount of coupling among agents in a team) may involve varying levels of interaction and autonomy, which brings different challenges with respect to the mode of human-aware planning required. For example, when the robot is providing direct assistance to its human teammate, it has to estimate the perceived usefulness of the intervention to the human's plan (plan level coordination). However, if the robot is tasked with its own goals but is operating in parallel to its human colleague, then beyond its own plan it also needs to be aware of the intentions of the human in order to anticipate and avoid conflicting interests (resource level coordination). Again, if the robot is trying to form a team with a human, it must evaluate how the team formation affects the agents involved beyond just joint plan optimality (goal-level coordination). In this task, we propose to modulate human-robot cooperation in human-robot teams at three different levels –

– Resource level: The robot modifies its planning process so as to respect the human's plans in terms of the resources required - this is a form of indirect assistance, the human does not explicitly realize the effect of the robot's deliberative process. We propose to develop different modes of behavior that can support such interactions, and investigate the pros and cons associated with them. (Chakraborti *et. al.,* 2016)

– Plan level: The robot provides assistance keeping in mind the perceived usefulness of that assistance from the human's point of view (e.g. in situations when the human is not anticipating the help.) Note that, unlike the previous case, this involves direct intervention to the human's plan. We plan to develop algorithms that can exploit opportunities for such assistance in the context of human-robot teams. (Chakraborti *et. al.,* 2015)

– Goal level: Even at the time of team formation, the robot must be able to reason about the implications of forming a team with a human, in terms of how it affects each other's intentions, and the commitments or expectations imposed thereof. Here, we investigate how a robot can pick what tasks to do while being aware of potential ad-hoc teaming opportunities with humans around it, on how the choice affects, and is affected by the human, and how this can be done with minimum prior coordination. (Chakraborti *et. al.,* 2016)



(a) Compromise.

(b) Opportunism.

(c) Negotiation.

Figure 8. Planning for conflict avoidance.

The Planner. Traditional planners provide no direct way to handle such complex interaction contraints within the planning process. Note here that since the execution of the plans of the agents is occurring in parallel, the uncertainty is evolving at the time of execution, and hence the uncertainty cannot be captured from the goal states of the recognized plans alone, and consequently cannot be simply compiled away to the initial state uncertainty for the robot and solved as a conformant plan. Similarly, the problem does not directly compile into action costs in a metric planning instance because the profiles themselves are varying with time. Thus we need a planner that can handle these resource constraints that are both stochastic and non-stationary due to the uncertainty in the environment. To this end we propose to use integer program or IP-based planner as an elegant way to model these complex interaction constraints during the plan generation process.

Figure 8 above shows some of the interesting behaviors of the robot produced by the planner in a typical USAR setting involving a human with a triage goal in room1 and the robot with an independent medkit delivery goal in the hallway. In 8(a) the robot *compromises* by getting the medkit further away from room3 because the human is expected to use medkit1 while in 8(b) the robot is *opportunistic* by anticipating favorable change in the world, and waits for the human to bring the medkit closer. Finally in 8(c) the robot decouples the conflict in the plans and *negotiates* a time of use before the human's. The policies of compromise and opportunism for the robot are complementary to negotiation in the event the latter fails. Thus, for example, the robot might choose to communicate a negotiation strategy to the human, but fall back on a compromise if that fails. Finally, as shown below in Figure 9, the robot can help him the commander by intercepting him with a medkit in the hallway so he need not fetch one himself – serendipity! This mode of interaction is not expected from the human's point of view and hence the planner must take into special considerations while producing such plans, given its teammate is not planning to exploit this.

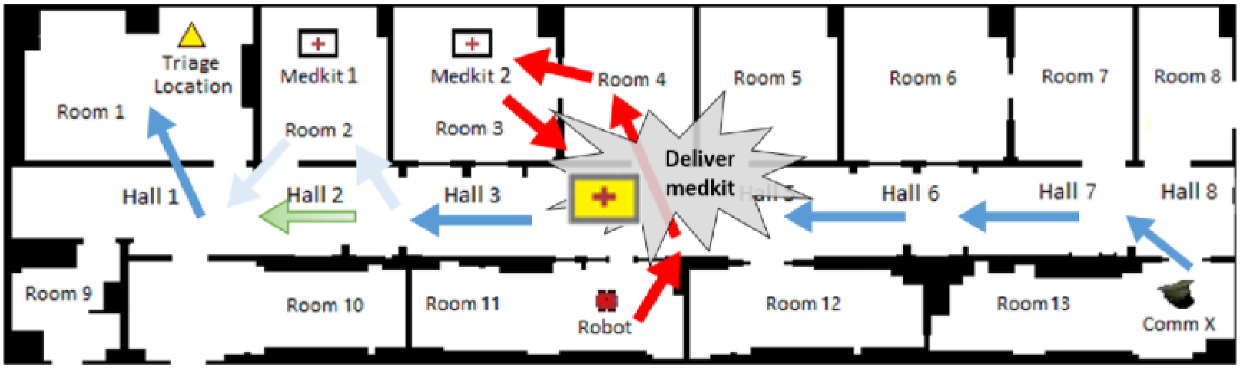
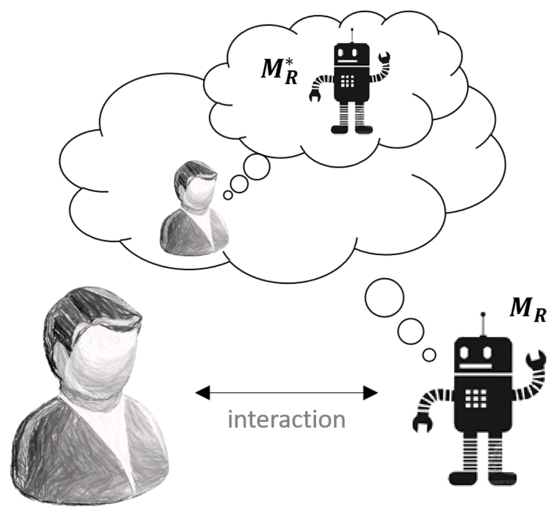


Figure 9. The robot intervenes in the sub-optimal human plan and produces an unexpected happy event (serendipity) with no communication.

Planning and Signaling Explainable Behavior

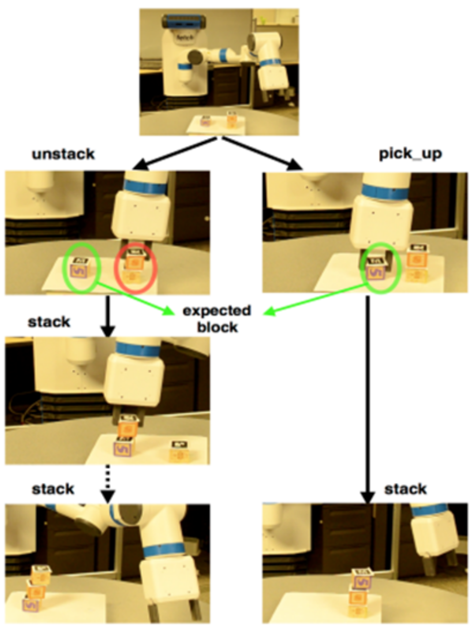


An important issue in cognitive teaming is for the humans to also maintain an understanding of the robots. Given that it is impractical to require humans to maintain complex models as the robots do, it requires the robots to either clearly signal their intents, or act in an explainable way such that their behaviors would be explicable and predictable to the humans.

To behave in an explainable way, the robot not only needs to construct behavior based on its own model (M_R), but must also consider the interpretation of robot behavior from the human perspective, which is influenced by the expected model of the robot (M_R^*). The technical challenge here is that the robot does not have direct access to M_R^* since it is deeply hidden from the human side. The robot would have to learn M_R^* and use it to modify its own plan generation. In this task, we plan to introduce plan explainability and predictability for such agents (Zhang et al., 2015) that can autonomously construct plans given goals. The consideration of both M_R and M_R^* enables robots to synthesize plans that do not surprise the humans, or affect them adversely. We propose to learn these measures (explainability and predictability) using machine learning techniques that can capture sequential information (e.g., CRFs (Lafferty, McCallum & Pereira, 2001) and LSTMs (Hochreiter & Schmidhuber, 1997)). The proposed measures will have a variety of applications (e.g., achieving fluent human-robot interaction and safety in human-robot teaming).

When explainable behavior alone is insufficient, (i.e., when a behavior is desirable but expected to be difficult for the humans to comprehend), signaling actions that involve explicit communication are required. In the first phase, we propose to employ cognitive cues (e.g., signaling actions) that are tightly integrated with the behaviors to project the intent of the robot so that it can be read directly by the humans. Our explainable plan generation method can automatically produce the set of signaling actions necessary for a given plan. In the second phase, more explicit communication modalities would be explored. For example, we plan to employ plan projection techniques to display visual information. Natural language communication between the human and robot would also be investigated with design informed by our human systems engineering expertise.

To compute the *explainability* and *predictability* measures, first, we postulate that humans understand robot plans by associating high level tasks with robot actions, which can be considered as a labeling process. We learn the labeling scheme of humans for robot plans from training examples using conditional random fields (CRFs). Then, we use the learned model to label a new plan to compute its explainability and predictability. These measures can be used by the robots to proactively choose plans, or can be directly incorporated into the planning process to generate plans that are more explainable and predictable.



Given a domain, the challenge is to find a plan for a given problem that satisfies the following:

$$arg\ min_{\pi_{M_R}}\ cost(\pi_{M_R}) + \alpha \cdot\ dist(\pi_{M_R}, \pi_{M_R^*})$$

where π_{M_R} is a plan that is constructed using M_R (i.e., the plan constructed by the agent), $\pi_{M_R^*}$ is a plan that is constructed using M_R^* (i.e., human's anticipation of the agent's plan), $cost$ returns the cost of a plan, $dist$ returns the distance between two plans, and α is the relative weight. The goal of Eq. 1 is to find a plan that minimizes a weighted sum of the cost of the agent plan and the distance (i.e., difference) between the two plans. Since the agent model M_R is often given, the challenge is on the second part in Eq. 1 since M_R^* is inherently hidden, difficult to convey, and can be arbitrarily different from M_R. Hence, we use a learning method to directly learn the $dist$ function. We postulate that $dist$ can be functionally decomposed as:

$$dist(\pi_{M_R}, \pi_{M_R^*}) = \mathrm{F} \circ \mathrm{L}^*(\pi_{M_R})$$

where F is a domain specific function that takes plan labels as input, and L^* is the labeling scheme of the human for agent plans based on M_R^*. As a result, Eq. 1 now becomes:

$$arg\ min_{\pi_{M_R}}\ cost(\pi_{M_R}) + \alpha \cdot\ \mathrm{F} \circ \mathrm{L}^*_{CRF}(\pi_{M_R} | \{S_i | S_i = L^*(\pi^i_{M_R})\})$$

where S_i is the set of training examples and L^*_{CRF} is the learned model of L^*.

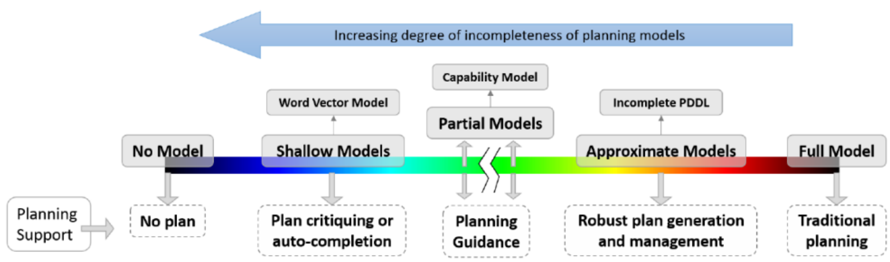
Given a domain, the explainability θ_π of an agent plan π is computed by a mapping, $F_\theta: \boldsymbol{L}_\pi \rightarrow [0,1]$ (with 1 being the most explainable). $\boldsymbol{L}_\pi$ above denotes the sequence of action labels for π. To compute θ_π and β_π for a given plan π, the challenge is to provide a label for each action. This requires us to learn the labeling scheme of humans (i.e., L^* in the equation above) from training examples and then apply the learned model to π (i.e., L^*_{CRF} in the equation above). To formulate a learning method, we consider the sequence of labels as hidden variables. The plan that is executed by the agent constitutes the observations.

The graphical model that we choose for the learning method is conditional random fields (CRFs) which have been shown to relax assumptions about the input and output sequence distributions, unlike HMMs, and hence are more flexible. The distributions that are captured by CRFs have the following form

$$p(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{Z}\,\Pi_A \Phi(\boldsymbol{x}_A,\boldsymbol{y}_A)$$

In the equations above, $\boldsymbol{x}$ represents the sequence of observations, $\boldsymbol{y}$ represents the sequence of hidden variables. $\Phi(\boldsymbol{x}_A,\boldsymbol{y}_A)$ represents a factor that is related to a subgraph in the CRF model associated with variables $\boldsymbol{x}_A$ and $\boldsymbol{y}_A$. In our context, $\boldsymbol{x}$ are the observations made during the execution of the action sequence of a plan; $\boldsymbol{y}$ are the action labels. Each factor is associated with a set of features observed during the plan execution. Given a set of training examples, we can train the CRF model. During plan synthesis (or testing), the agent needs to synthesize a plan that is explainable and predictable. We propose to investigate two ways to use the learned CRF model (Zhang *et. al.,* 2016) as plan selection and plan synthesis.

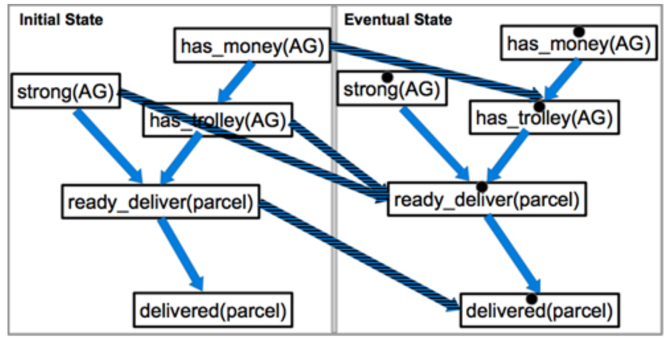
<u>Representation and Reasoning with Learnable Incomplete Models</u>



Because it is often impractical for human models to be provided in a hand-coded form, we need to look for models that can be easily learned and refined from traces of the observed human behavior. Such learned models will be, of necessity, incomplete, and that there can be a spectrum of such incomplete models (Kambhampati, 2007). In this task, we propose to develop and evaluate two representative points in that spectrum: Capability Models, which are partial models, and Action Vector models which are shallow models. These are especially relevant for human-robot interactions given the robot's partial model of the human. An important motivation for exploring these models is that the training data requirements for the partial and shallow models are significantly lower than the requirements for traditional PDDL models. With these models, we can then develop algorithms for robots to plan interactions in human-robot teaming (Chakraborti et al., 2015; Chakraborti, Zhang, Smith & Kambhampati, 2016).

<u>Capability Model</u>

Capability model (Zhang, Sreedharan & Kambhampati, 2015) is one of the learnable (partial) models we propose to develop. The underlying structure of a capability model is a generalized 2-slice temporal Bayes network (2-TBN). In this task, we propose to investigate the representation of this model and its applicability in human modeling. In particular, we will first propose efficient methods to learn these models, and then explore the relationship between Bayesian inference on capability models with planning and plan recognition. Compared to MDP and action based models, which do not naturally represent partial and incomplete models, capability models have their unique benefits and limitations. In this aspect, capability models should not be considered as a competitor to these more complete models for automated planning. Instead, it is useful when more complete models are difficult to obtain or only partial information can be collected to learn the models. This is especially suitable for human modeling.



The representation of a capability model is a generalization of a two time slice dynamic Bayesian network (2-TBN). Each node at the first time slice (also called a *fact node*) represents a variable that is used in the specification of initial world states. Furthermore, for each fact node, there is also a corresponding node in the second time slice, which represents the eventual state of the fact node after an operation. These corresponding nodes are called eventual nodes or e-nodes. The state specified by the fact nodes is referred to as the *initial state*, and the state specified by the e-nodes is referred to as the *eventual state*. The links between the e-nodes, as well as links from the fact nodes to e-nodes in the model, follow the same causal relationships between the fact nodes. Furthermore, there is also a link from each fact node to its corresponding e-node. Figure \ref{fig:cap-model} presents a capability model.

Capability. Given an agent ϕ, a capability specified as a mapping $S_\phi \times S_\phi \rightarrow [0,1]$, is an assertion about the probability of the existence of plans (for ϕ) that connect an initial state (i.e., the first component on the left hand side of the mapping) to an eventual state (i.e., the second component). A capability is also denoted as $s_I \Rightarrow s_E$ (i.e., the initial state $\Rightarrow$ the eventual state) when we do not need to reference the associated probability value. The probability value is denoted by $P(s_I \Rightarrow s_E)$. First, both s_I and s_E can be partial states. We refer to a plan that satisfies the specifications of a capability as an *operation*.

Given the partial nature of the model, the plan generated would also be incomplete. Hence, it would be interesting to investigate how to use these incomplete plans to guide planning and decision making. We also propose to establish the relationships between capability models and traditional complete models. In terms of applications, we will explore using capability models to model human behavior in various human-robot teaming scenarios such as urban search and rescue, and household assistance.

Action Vector Model

Action vector model (Tian, Zhuo & Kambhampati, 2015) represents the class of shallow models, which are learned from just short plan fragments, and essentially captures the affinity between actions based on how often they occur in specific contexts. While we propose to focus initially on a one-shot recognition task (Tian, Zhuo & Kambhampati, 2015), in practice, human-in-the-loop planning will consist of multiple iterations and long-term tasks, with our model recognizing the plan and suggesting action addition alternatives; the human making a selection and revising the plan. It is interesting to see how action vector models can complement the other models when even learning partial models (e.g., capability models) becomes a challenge.

We assume as input a set of plan fragments, where each plan fragment is just a sequence of actions. Since actions are denoted by a name strings, actions can be viewed as words, and a plan can be viewed as a sentence. Furthermore, the plan library $\mathcal{L}$ can be seen as a corpus, and the set of all possible actions a is the vocabulary. We thus can learn the vector representations for actions using the Skip-gram model (Mikolov *et. al.* 2013) with hierarchical softmax, which has been shown an efficient method for learning high-quality vector representations of words from unstructured corpora (Tian *et. al.* 2016). The objective of the Skip-gram model is to learn vector representations for predicting the surrounding words in a sentence or document. Given a corpus $\mathcal{C}$, composed of a sequence of training words $\langle w_1, w_2, \ldots, w_T \rangle$, where $T = |\mathcal{C}|$, the Skip-gram model maximizes the average log probability as follows –

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0} \log\left(p\left(w_{\{t+j\}} \middle| w_t\right)\right)$$

where c is the size of the training window or context. The basic probability $p\left(w_{t+j} \mid w_t\right)$ is defined by the hierarchical softmax, which uses a binary tree representation of the output layer with the K words as its leaves and for each node, explicitly represents the relative probabilities of its child nodes. For each leaf node, there is a unique path from the root to the node, and this path is used to estimate the probability of the word represented by the leaf node. Each inner node has an output vector $v'_{n(w,j)}$, and the probability of a word being the output is defined by

$$p\left(w_{t+j} \mid w_t\right) = \prod_{i=1}^{L(w_{t+j})-1}\left\{\sigma\left(\mathbb{I}\left(n\left(w_{t+j}, i+1\right) = child\left(n\left(w_{t+j}, i\right)\right)\right) \cdot v_{n\left(w_{t+j}, i\right)} \cdot v_{w_t}\right)\right\},$$

where $(x) = \frac{1}{1+e^{-x}}$. $L(w)$ is the length from the root to the word w in the binary tree, e.g., $L(w) = 4$ if there are four nodes from the root to w. $n(w, i)$ is the i^{th} node from the root to w, e.g., $n(w, 1) = root$ and $n(w, L(w)) = w$. $child(n)$ is a fixed child (e.g., left child) of node n. v_n is the vector representation of the inner node n. v_{w_t} is the input vector representation of word w_t. The identity function $\mathbb{I}(x)$ is 1 if x is true; otherwise it is -1.

We can thus build vector representations of actions by maximizing the equation above, with corpora or plan libraries $\mathcal{L}$ as input. We will exploit the vector representations to predict possible plan $\tilde{\pi}$ of the human-teammate based on a robot's experiences of interaction.

9.0 Technical Approach: Human Systems Engineering

The purpose of the human systems engineering studies is to test, in an iterative test-re-design process, the robot algorithms developed. Specifically, the objective is:

Conduct human systems engineering studies to evaluate and iterate on the robot representations and algorithms and to inform future development.

The tasks needed to achieve this objective are as follows:

a. Implement team behaviors
b. Test in reconnaissance scenario with one external human one or two internal cognitive robots
c. Evaluate and iterate

This set of studies will involve incorporation of actual robot algorithms inspired by previous human studies and our deepened understanding of effective teammate behavior. These studies will not only provide a test of the robot teammate and human-robot teaming, but will also serve as a test of our understanding of teamwork and team cognition in this task setting. Iterations will continue until the test is conducted with a fully functioning robot-human team and compared to an all-human team.

For this study, a simulated environment will be built either within Webots (https://www.cyberbotics.com), or Microsoft Malmo platform (https://www.microsoft.com/en-us/research/project/project-malmo/) to replicate the Minecraft scenario. Thirty individuals will be tested in a human-in-the-loop study. The same metrics of team effectiveness, team SA, workload and trust will be applied, focusing on the human and comparison of metrics to previous studies involving all humans.

The following research questions will be addressed:

- Is the effectiveness and SA of human-robot teams comparable to human-human teams?
- How do explicable behavior and excuses influence effectiveness, SA, and trust?

10.0 Technical Approach: Long-term Teaming and RPAs

Option years will involve extending the results across time and task domain. The objective addressed here is to:

Extend findings to Remotely Piloted Aircraft System environment and to longer term relationships.

The tasks needed to achieve this objective are as follows:

a. Extend lessons from this research to the development of approaches that will support effective manned-unmanned aircraft teaming (e.g. A "Loyal Wingman: RPA operating as assistant to an F-35 or F-22).
b. In reconnaissance domain bring humans back to work with same robots and examine dynamics of trust and shared SA over time

Remotely Piloted Aircraft Systems

Although we expect that the robot teammate behaviors that lead to effective and situationally aware teams in the search and triage domain should generalize to human-robot teaming on other tasks, we will take this opportunity to test the replicability of the findings in a Remotely Piloted Aircraft (RPA) setting. Leveraging an existing synthetic task environment for RPAs (Figure 10) at ASU we can examine human-robot interaction in that context. Specifically, the CERTT-II task environment consists of multiple 40-minute missions wherein reconnaissance photographs of certain target waypoints must be obtained by three heterogeneous teammates: 1) Air Vehicle Operator (AVO or pilot) – controls the UAS's heading, altitude, and airspeed; 2) Data Exploitation, Mission Planning, and Communications (DEMPC or navigator) – generates a dynamic flight plan and issues speed and altitude restrictions; and 3) Payload Operator (PLO or sensor operator) – monitors sensor equipment, negotiates with the AVO on speed and altitude in order to take a good photo of the target waypoints. Communication within the three-agent UAS teams occurred over a text-based communications system. In order to take a good photo for each target waypoint, the coordination among the three team members needed to follow an optimal coordination sequence:

Figure 10. CERTT RPA testbed.

***Information-Negotiation-Feedback (INF)*:** the navigator (DEMPC) provides the *information* about the upcoming target waypoint to the pilot (AVO); the AVO *negotiates* with the photographer (PLO) about an appropriate altitude and airspeed for the target waypoints and required camera settings; and finally, the PLO sends *feedback* to the AVO and the DEMPC about whether they have a good photo or not for the current target waypoint (Cooke, Gorman, Duran, & Taylor, 2007).

Due to the time-sensitive nature of the task, adherence to the INF coordination sequence is vital in order to maintain stable communication within the group and avoid any communication failures that would adversely affect team performance. Note that there are many possible temporal patterns of these coordination elements that maintain the same sequencing.

The cognitive robot in this task plays the role of the AVO or pilot. We have conducted studies with the synthetic teammate in the loop working with two humans (Figure 11). Communication occurs via text chat. In that environment we can examine whether the teammate behaviors discovered to lead to effective human-robot teaming in the search and triage domain translate to human-robot teaming in the RPA domain. Note that the synthetic teammate's communication capabilities are similarly limited making the findings from the search and triage domain not only relevant to the RPA domain, but potentially of value in increasing its ability to meaningfully interact with humans.

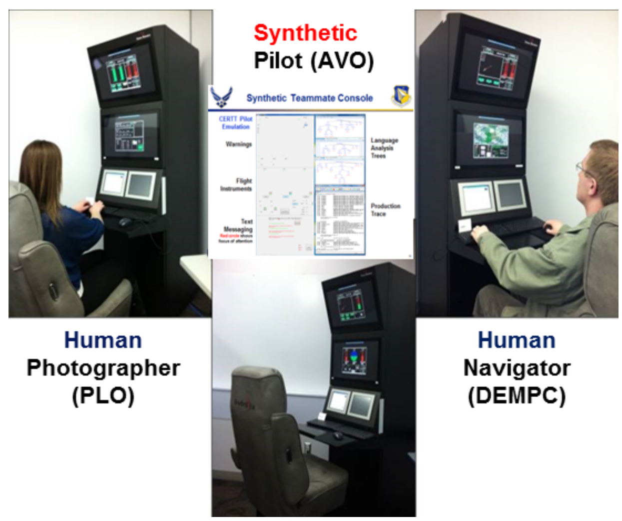


Figure 11. Two human operators interacting with the synthetic Air Vehicle Operator via text chat.

Long Term Teaming and Adaptive Trust

Many human-robot teaming tasks are not only complex, but can also span multiple episodes for an extended period of time. In such scenarios, the system's performance is dependent on how the teams perform in the current task, as well as how they perform in future tasks. A prerequisite to consider such long-term teaming is thus to maintain mental models of the agents that influence their interactions, and analyze how these models dynamically affect the teaming performance and how they evolve over time. A relevant issue that we are particularly interested in is how the trust between agents evolves during the long-term teaming and how autonomous agents can manipulate trust under different situations.

In this study we will reuse the test beds used in the human systems engineering study, but we will test human participants over a longer period of time. Specifically, ten participants will each interact with the robot to accomplish multiple missions over 10 sessions that span two months. Team effectiveness, team situation awareness, workload and trust will be measured with a special emphasis on how these factors change over time. The robot can also be manipulated to fail or to have inexplicable behavior early or late in the series of missions and the effect of the timing of this failure will also be examined.

Results from this study will provide a deeper look at the evolution of trust over time that is more realistic than what is typically measured in a one session study. Findings will also speak to how varying levels of trust impacts team effectiveness and situation awareness.

11.0 Significance of Contribution

The work outlined in this proposal will contribute to our understanding of teaming effectiveness and especially characteristics of individuals that are tied to effective teamwork, team situation awareness, trust and workload. Furthermore the efforts will advance the state of the art in robot models that comprehend human intent and at the same time exhibit behavior that can be understood by humans. In addition, we expect to make significant headway on problems of human-robot interaction through limited dialog, long-term human-robot teaming and the evolution of trust, and generalizability of findings across tasks. Overall this research should contribute to the Air Force's goal of improved human-autonomy teaming.

The proposed research is timely as it resonates with the increased interest in human-aware AI systems. The recent White House report (2016) on "*The National AI R&D Strategic Plan"* recommends developing effective methods for Human-AI collaboration as an important strategic direction, and recommends seeking new algorithms for human-aware AI. The proposed research is also in broad concordance with the AFOSR report on "*Recommendations for Research on Trust in Autonomy"* (Gratch, Friedland & Knott, 2016). In particular, we consider how to make the machines (robots) develop and maintain mental models of the human collaborators, and how to make them behave in such a way as to adapt to the human user's expectations, and, in long term, engender trust in the users.

12.0 References

Admoni, H., Weng, T., Hayes, B. and Scassellati. (2016) Robot Nonverbal Behavior Improves Task Performance In Difficult Collaborations. HRI 2016: 51-58.

Bartlett, C. E., & Cooke, N. J. (2015). Human-Robot Teaming in Urban Search and Rescue. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 59, No. 1, pp. 250-254). SAGE Publications.

Bartlett, C., McNeese, N. J., Cooke, N. J., Zhang, Y., & Kambhampati, S. (under revision). Human Robot Teaming in Urban Search and Rescue.

Bolstad, C. A., & Endsley, M. R. (1999). Shared mental models and shared displays: An empirical evaluation of team performance. In *proceedings of the human factors and ergonomics society annual meeting* (Vol. 43, No. 3, pp. 213-217). SAGE Publications.

Bolstad, C.A. and Endsley, M.R., 2003, Measuring shared and team situation awareness in the Army's future objective force. In Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, pp. 369–373 (Santa Monica, CA: Human Factors and Ergonomics Society).

Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. Journal of Organizational Behavior, 22, 195–202. doi: 10.1002/job.82

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In J. Castellan, Jr. (Ed.), Current issues in individual and group decision making (pp. 221-246). Hillsdale, NJ: Erlbaum.

Casper, J., & Murphy, R. R. (2003). Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *33*(3), 367-385.

Castro, A., Admoni, H. Scassellati, B. Effects of form and motion on judgments of social robots' animacy, likability, trustworthiness and unpleasantness. Int. J. Hum.-Comput. Stud. 90: 27-38 (2016)

Chiou, E. K., & Lee, J. D. (2016). Cooperation in human-agent systems to support resilience: A microworld experiment. *Human Factors, 58*, 846-863.

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive Science*, *37*(2), 255-285.

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, *49*(12-13), 1312-1325.

Chakraborti, T., Briggs, G., Talamadupula, K., Zhang, Y., Scheutz, M., Smith, D., & Kambhampati, S. (2015, September). Planning for serendipity. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on* (pp. 5300-5306). IEEE.

Chakraborti, T., Talamadupula, K., Zhang, Y., & Kambhampati, S. (2016) A Formal Framework for Studying Interaction in Human-Robot Societies. *AAAI 2016 Workshop on Symbiotic Cognitive Systems.*

Chakraborti, T., Zhang, Y., Smith, D. E., & Kambhampati, S. (2016). Planning with Resource Conflicts in Human-Robot Cohabitation. In *Proc. of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Connors, E. S., Strater, L. D., Riley, J. M., & Endsley, M. R. (2008). Applying SA-Oriented Design to the Integrative Collaborative Control of Multiple Unmanned Vehicle Systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, No. 4, pp. 211-215). SAGE Publications.

Dragan, A., Lee, K. and Srinivasa, S. Legibility and predictability of robot motion. HRI 2013: 301-308

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 32, No. 2, pp. 97-101). SAGE Publications.

Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 4-32.

Endsley, M., & Jones, W. M. (1997). *Situation Awareness Information Dominance & Information Warfare*. LOGICON TECHNICAL SERVICES INC DAYTON OH.

Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. Human Factors, 41, 312–325.

Fiore, S. M., Salas, E., Cuevas, H. M., & Bowers, C. A. (2003). Distributed coordination space: toward a theory of distributed team process and performance. *Theoretical Issues in Ergonomics Science*, *4*(3-4), 340-364.

Fong, T.W., Nourbakshsh, I., Dautenhahn, K. (2003) A survey of socially interactive robots. Robotics and Autonoumous Systems.

Geffner, H. and Bonet, B. (2013) A Concise Introduction to Models and Methods for Automated Planning. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers.

Ghallab, M., Nau, D. and Traverso, P. (2004) Automated planning: theory \& practice. Access Online via Elsevier, 2004.

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, *49*(12-13), 1312-1325.

Grigore, E., Pereira, A., Zhou, I., Wang, D. and Scassellati, B. (2016) Talk to Me: Verbal Communication Improves Perceptions of Friendship and Social Presence in Human-Robot Interaction. IVA 2016: 51-63

Gratch, J., Friedland, P., Knott, B. (2016). Recommendations for Research on Trust in Autonomy. In 5th International Workshop on Human-Agent Interaction Design & Models . New York, NY.

Hayes, B. and Scassellati, B. (2016) Autonomously constructing hierarchical task networks for planning and human-robot collaboration. ICRA 2016: 5469-5476

Hayes, B. and Scassellati, B. (2015) Effective robot teammate behaviors for supporting sequential manipulation tasks. IROS 2015: 6374-6380

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Kambhampati, S. (1997) Refinement planning as a unifying framework for plan synthesis.AI Magazine, 18(2):67—97.

Kambhampati, S. (2007). Model-lite planning for the web age masses: The challenges of planning with incomplete and evolving domain models. In *Proceedings of the National*

Conference on Artificial Intelligence (Vol. 22, No. 2, p. 1601). Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

Kambhampati, S. and Talamadupula, T. (2014). Human-in-the-Loop Planning and Decision Support. AAAI Tutorial. *rakaposhi.eas.asu.edu/hilp-tutorial*

Kapalo, K. A., Phillips, E., & Fiore, S. M. (2016). The Application and Extension of the Human-Animal Team Model to Better Understand Human-Robot Interaction: Recommendations for Further Research. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 60, No. 1, pp. 1225-1229). SAGE Publications.

Kruijff, G., Colas, F., Svoboda, T., Van Diggelen, J., Balmer, P., Pirri, F., & Worst, R. (2012). Designing intelligent robots for human-robot teaming in urban search & rescue. In *AAAI 2012 Spring Symposium on Designing Intelligent Robots*.

Kruijff, G. J. M., Janíček, M., Keshavdas, S., Larochelle, B., Zender, H., Smets, N. J., & Liu, M. (2014). Experience in system design for human-robot teaming in urban search and rescue. In *Field and Service Robotics* (pp. 111-125). Springer Berlin Heidelberg.

Kruse, T., Pandey, A.K., Alami, R. and Kirsch, A. (2013) Human-aware robot navigation: A survey. Robot. Auton. Syst., 61(12):1726—1743.

Lafferty, J., McCallum, A., & Pereira, F. C. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *Proceedings of the Eighteenth International Conference on Machine Learning (ICML)*. 282–289.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of applied psychology*, *85*(2), 273.

McNeese, N. J., & Reddy, M. C. (2015). Articulating and Understanding the Development of a Team Mental Model in a Distributed Medium. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 59, No. 1, pp. 240-244). SAGE Publications.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., and Dean, J. (2013) Distributed representations of words and phrasesand their compositionality. In NIPS, pages 3111–3119.

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management*.

Montague, E., Asan, O., & Chiou, E. (2013). Organizational and technological correlates of nurses' trust in a smart intravenous pump. *Computers, Informatics, Nursing*, *31*(3), 142–149.

Montague, E., & Chiou, E. (2014). Trust in complex work systems: A focus on information and communication technologies. In C. Korunka & P. Hoonakker (Eds.), *The Impact of ICT on Quality of Working Life* (pp. 143–152). Dordrecht: Springer.

Narayanan, V., Zhang, Y., Mendoza, N., & Kambhampati, S. (2015). Automated Planning for Peer-to-peer Teaming and its Evaluation in Remote Human-Robot Interaction. In *HRI (Extended Abstracts)* (pp. 161-162).

O'Leary, M. B., & Cummings, J. N. (2007). The spatial, temporal, and configurational characteristics of geographic dispersion in teams. *MIS quarterly*, *31*(3), 433-452.

Parasuraman, R., Barnes, M., Cosenzo, K., & Mulgund, S. (2007). *Adaptive automation for human-robot teaming in future command and control systems*. ARMY RESEARCH LAB

ABERDEEN PROVING GROUND MD HUMAN RESEARCH AND ENGINEERING DIRECTORATE.

Riley, J. M., & Endsley, M. R. (2004). The hunt for situation awareness: Human-robot interaction in search and rescue. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 48, No. 3, pp. 693-697). SAGE Publications.

Riley, J. M., & Endsley, M. R. (2005, September). Situation awareness in HRI with collaborating remotely piloted vehicles. In *proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 49, No. 3, pp. 407-411). SAGE Publications.

Riley, J. M., Strater, L. D., Sethumadhavan, A., Davis, F., Tharanathan, A., & Kokini, C. (2008). Performance and situation awareness effects in collaborative robot control with automation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, No. 4, pp. 242-246). SAGE Publications.

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 540-547.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), Teams: Their training and performance (pp.3–29). Norwood, NJ: Ablex.

Shah, J., Wiken, J., Williams, B. and Breazeal, C. (2011) Improved human-robot team performance using chaski, a human-inspired plan execution system. Proceedings of the 6th international conference on Human-robot interaction}, pages 29--36. ACM.

Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 137-148.

Stagl, K. C., Salas, E., Rosen, M. A., Priest, H. A., Burke, C. S., Goodwin, G. F., & Johnston, J. H. (2007). Distributed team performance: A multi-level review of distribution, demography, and decision making. *Research in multi-level issues*, *6*, 11-61.

Talamadupula, K., Benton, J., Kambhampati, S., Schermerhorn, P., & Scheutz, M. (2010). Planning for human-robot teaming in open worlds. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *1*(2), 14.

Tian, X., Zhuo, H. H., & Kambhampati, S. (2015). Discovering Underlying Plans Based on Distributed Representations of Actions. *In International Conference on Autonomous Agents and Multiagent Systems (AAMAS).*

United States Air Force. (2015). Autonomous Horizons: System Autonomy in the Air Force – A Path to the Future (Volume I: Human-Autonomy Teaming).

Unhelkar, V.V., Siu, Ho-Chit and Shah, J. (2014) Comparative performance of human and mobile robotic assistants in collaborative fetch-and-deliver tasks. Proceedings of the 2014 ACM/IEEE International Conference on Human-robot Interaction}, HRI '14, pages 82--89, New York, NY, USA.

Wiltshire, T. J., Barber, D., & Fiore, S. M. (2013). Towards modeling social-cognitive mechanisms in robots to facilitate human-robot teaming. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 57, No. 1, pp. 1278-1282). SAGE Publications.

Zhang, Y., Sreedharan, S., & Kambhampati, S. (2015). Capability models and their applications in planning. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (pp. 1151-1159). International Foundation for Autonomous Agents and Multiagent Systems.

Zhang, Y., Narayanan, V., Chakraborti, T., & Kambhampati, S. (2015). A human factors analysis of proactive support in human-robot teaming. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on* (pp. 3586-3593). IEEE.

Zhang, Y., Sreedharan, S., Kulkarni, A., Chakrabarti, T., Zhou, H., Kambhampati, S., (2016) (Under Submission) Plan Explainability and Predictability for Robot Task Planning. Technical report.

13.0 Personnel

Nancy J. Cooke will serve as PI and Subbarao Kambhampati as Co-PI. Nathan McNeese, Erin Chiou, and Yu Zhang will serve as Co-Is. Mica Endsley will serve as a consultant to the project. Kambhampati and Zhang will oversee the robotics development and Cooke, Chiou, and McNeese will oversee the teaming studies. Cooke will plan technical exchange meetings to facilitate coordination among the team members.

Nancy J. Cooke

Nancy J. Cooke is a professor of Human Systems Engineering at Arizona State University and is Science Director of the Cognitive Engineering Research Institute in Mesa, AZ. She received her PhD in Cognitive Psychology from New Mexico State University in 1987. Dr. Cooke is currently President of the Human Factors and Ergonomics Society and a National Associate of The National Academies of Sciences, Engineering, and Medicine. She also recently chaired a study panel for the National Academies on the Enhancing the Effectiveness of Team Science. Dr. Cooke was a member of the US Air Force Scientific Advisory board from 2008-2012. In 2014 Dr. Cooke received the Human Factors and Ergonomics Society's Arnold M. Small President's Distinguished Service Award for career-long contributions to human factors. Dr. Cooke's research interests include the study of individual and team cognition and its application to the development of cognitive and knowledge engineering methodologies, human-robot teaming, sensor operator threat detection, cyber security, intelligence analysis, remotely piloted aircraft systems, human-robot interaction, healthcare systems, and emergency response systems. Dr. Cooke has published extensively on topics of team cognition and knowledge elicitation.

Select Publications

Peer Reviewed Journal Articles

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies, 41,* 801-849.

Gorman, J.C., Cooke, N. J., & Winner, J.L. (2006). Measuring team situation awareness in decentralized command and control systems. *Ergonomics, 49,* 1312-1325.

Gorman, J. C., Cooke, N. J., & Amazeen, P. G. (2010). Training adaptive teams. *Human Factors, 52,* 295-307. [Winner of HFES 2010 Jerome Ely Award]

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J.L. (2013). Interactive Team Cognition, *Cognitive Science, 37, 255-285,* DOI: 10.1111/cogs.12009.

Cooke, N. J. (2015). Team cognition as interaction. *Current Directions in Psychological Science, 34,* 415-419.

Books & Book Chapters

Cooke, N. J., Stout, R.J., & Salas, E. (2001). A knowledge elicitation approach to the measurement of team situation awareness. In M. McNeese, M. Endsley, & E. Salas, (Eds.), *New Trends in Cooperative Activities: System Dynamics in Complex Settings,* pp. 114-139. Santa Monica, CA: Human Factors.

Cooke, N. J., Pringle, H., Pedersen, H., & Connor, O. (Eds.) (2006*). Human Factors of Remotely Operated Vehicles.* Volume in *Advances in Human Performance and Cognitive Engineering Research* Series, Elsevier.

Cooke, N. J. & Durso, F. (2008). *Stories of Modern Technology Failures and Cognitive Engineering Successes*, Taylor and Francis.

Cooke, N. J., Rowe, L. R. Bennett, W., & Joralmon, D. Q. (in press). *Remotely Piloted Aircraft Systems: A Human Systems Integration Perspective.* Wiley

US Air Force Scientific Advisory Board Reports *(written by committee)*

USAF SAB. *Virtual Training Technologies*. (2009). Report of the USAF Scientific Advisory Board.

USAF SAB. *Operating Next-Generation Remotely Piloted Aircraft for Irregular Warfare.* (2010). Report of the USAF Scientific Advisory Board.

USAF SAB. *Sensor Data Exploitation*. (2011). Report of the USAF Scientific Advisory Board.

USAF SAB. *Cyber Situation Awareness.* (2012). Report of the USAF Scientific Advisory Board.

Subbarao Kambhampati

Subbarao Kambhampati is a professor of Computer Science & Engineering at Arizona State University, where he directs the Yochan research group. He received his Ph.D. in Computer Science from University of Maryland in 1989. Kambhampati is currently the President of the Association for the Advancement of Artificial Intelligence (AAAI). He is also an elected fellow of AAAI. He is a Trustee of the International Joint Conference on Artificial Intelligence (IJCAI Inc.), and was the program chair for IJCAI 2016. He also co-chaired AAAI 2005, AIPS 2000 and ICAPS 2013. Kambhampati's research interests are primarily in Artificial Intelligence, and include planning and decision making, human-robot teaming and human-aware AI. He also works in the area of social media analysis and data integration. Kambhampati has published extensively and is well-cited; Google Scholar counts over 7600 citations to his work, and gives an *h-index* of 45. His work has been supported by many competitive research grants, including an NSF Young Investigator award (1994), grants from ONR, ARO and DARPA, as well as IBM and Google. Papers from his group have been honored with influential paper awards (ICAPS 2010), as well as best paper awards and nominations (WWW 2010, ICAPS 2014, and AAMAS 2016). Two of his students received best dissertation recognition in the planning community (ICAPS 2007). Kambhampati is acknowledged as an authority on automated planning and several of his contributions have found their way into popular AI textbooks. He received multiple awards for his teaching, including best teacher awards at college and department level, and a last lecture selection at the university level.

Select Publications

Kartik Talamadupula, J. Benton, Subbarao Kambhampati, Paul Schermerhorn, and Matthias Scheutz. Planning for Human-Robot Teaming in Open Worlds. ACM Transactions on Intelligent Systems and Technology. Vol 1. No. 2. 2010.

Tuan Ngueyn, Sarath Sreedharan and Subbarao Kambhampati. Robust Planning with Incomplete Domain Models. *Artificial Intelligence. 2016 (To appear)*

Nan Li, William Cushing, Subbarao Kambhampati and Sungwook Yoon. Learning Probabilistic Hierarchical Task Networks to Capture User Planning Preferences. ACM Transactions on Intelligent Systems and Technology. ACM TIST 5(2): 29 (2014)

Tuan Nguyen, Minh Do, Alfonso Gerevini, Ivan Serina, Biplav Srivastava and Subbara Kambhampati. Planning with partial preference models Artificial Intelligence. Vol 190. Pages 1-31. Oct 2012.

Xin Tian, Hankz Hankui Zhuo and Subbarao Kambhampati. Discovering Underlying Plans Based on Distributed Representations of Actions. AAMAS 2016 (Nominated for AAMAS 2016 Best Student Paper Award)

Yu Zhang, Sarath Sreedharan & Subbarao Kambhampati A Formal Analysis of Required Cooperation in Multi-agent Planning ICAPS 2016.

Tathagata Chakraborti, Yu Zhang, Subbarao Kambhampati. Planning with Resource Con- flicts in Human-Robot Cohabitation AAMAS 2016.

Tathagata Chakraborti, Gordon Briggs, Kartik Talamadupula, Yu Zhang, Matthias Scheutz, David Smith and Subbarao Kambhampati Planning for Serendipity. IROS 2015

Yu (Tony) Zhang, Vignesh Narayanan, Tathagata Chakraborty & Subbarao Kambhampati. A Human Factors Analysis of Proactive Assistance in Human-robot Teaming. IROS 2015.

Vignesh Narayanan, Yu Zhang, Nathaniel Mendoza and Subbarao Kambhampati. Automated Planning for Peer-to-peer Teaming and its Evaluation in Remote Human-Robot Interaction 10th ACM/IEEE Intl. Conf on Human Robot Interaction (HRI), 2015.

Yu Zhang, Sarath Sreedharan and Subbarao Kambhampati. Capability Models and their application in Multi-agent planning. AAMAS 2015.

Nathan J. McNeese

Nathan McNeese received a PhD in Information Sciences & Technology with a focus on Team Decision Making and Cognition from The Pennsylvania State University in the fall of 2014. Upon completion, he began an appointment with Arizona State University as a Postdoctoral Research Associate working in direct collaboration with Dr. Nancy Cooke. Prior to his PhD, he completed a B.S. in Psychology with a minor in Security Risk Analysis from The Pennsylvania State University. For over 10 years, Dr. McNeese has conducted research on many topics within a variety of different contexts. His current research interests span across the study of individual and team cognition, the development and design of human-centered tools and systems, continued development and refinement of cognitive engineering methods, the interplay between team interaction and team cognition, the development and application of remotely piloted aircraft systems, understanding teamwork in healthcare, and human robot interaction. Specific contexts that Dr. McNeese conducts his research in are homeland security, emergency crisis management, educational learning, sports, and healthcare systems. Throughout all of his research he uses knowledge elicitation methods to understand the context, work, and roles specific to the development of individual and team cognition. As a result of Dr. McNeese's work, his research has been published multiple times in numerous research communities.

Select Publications

McNeese, N. & Reddy, M. (2015) The Role of Team Cognition in Collaborative Information Seeking. *Journal of the Association for Information Science and Technology.* doi: 10.1002/asi.23614

McNeese, N., Cooke, N., Branaghan, R., Knobloch, A., & Taylor, A. (Accepted) Leveraging Expertise in Identification of the Emplacement of Improvised Explosive Devices for Mission Payload Operator Training. *Applied Ergonomics.*

McNeese, N., Khera, N., Wordingham, S., Arring, N., Nyquist, S., Gentry, A., Tomlinson, B., Cooke, N., & Sen, A. (2016). Applying the Science of Teams to Improve Critically Ill Cancer Care Delivery in Hematopoietic Stem Cell Transplant Patients. *Journal of Oncology Practice.*

McNeese, N. & Cooke, N. (2016). Team Cognition as a Mechanism for Developing Collaborative and Proactive Decision Support in Unmanned Aerial Systems. *18th International Conference on Human-Computer Interaction*. Toronto, CA.

Demir, M., McNeese, N., & Cooke, N. (2016). Team Communication Behaviors of The Human-Automation Teaming. *2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (COGSIMA)*. San Diego, CA. [Best Paper of 2016 Cogsima Conference]

McNeese, N., Reddy, M., & Friedenberg, E. (2014). Team Mental Models within Collaborative Information Seeking. *2014 Annual Meeting of the Human Factors and Ergonomic Society.* Chicago, IL. Human Factors and Ergonomics Society. October 27-31, 2014, pp. 335-339.

McNeese, N., & Reddy, M. (2015). Articulating and Understanding the Development of a Team Mental Model in a Distributed Medium. *2015 Annual Meeting of Human Factors and Ergonomic Society.* Los Angeles, CA. Human Factors and Ergonomics Society. October 26-30, 2015. pp. 240-44.

Demir, M., McNeese, N., Cooke, N., Ball, J, Myers, C. (2015). Synthetic Teammate Communication and Coordination with Humans. *2015 Annual Meeting of Human Factors and Ergonomic Society.* Los Angeles, CA. Human Factors and Ergonomics Society. October 26-30, 2015. pp. 951-955.

Yu Zhang

Yu Zhang is an Assistant Professor of Computer Science and Engineering at Arizona State University. He received the B.S. degree in software engineering from Huazhong University of Science and Technology, Wuhan, China, in 2006 and the M.S. and Ph.D. degrees in computer science from the University of Tennessee, Knoxville (UTK), in 2009 and 2012, respectively. He was a recipient of the University of Tennessee Chancellor's Citation Award for Extraordinary Professional Promise in 2012 and the University of Tennessee Graduate Student Fellowship in 2011. Dr. Zhang joined the Department of Computer Science and Engineering at Arizona State University as a Postdoctoral Researcher in July 2013, became a Research Assistant Professor in August 2015 and an Assistant Professor in 2016. His current research interests span the theories and practices of distributed robot systems, focusing on the coordination of heterogeneous robotic systems, and human-robot teaming systems, focusing on the enabling technologies for fluent teaming. His research work is also widely applicable to other domains such as proactive decision support systems and multi-agent systems. Dr. Zhang published extensively on topics in both AI and Robotics.

Select Publications

Peer Reviewed Journal Articles

Zhang, Y. and Parker, L. E. (2013). IQ-ASyMTRe: Forming Executable Coalitions for Tightly-Coupled Multirobot Tasks. *IEEE Transactions on Robotics*. Vol. 29. No. 2.

Zhang, Y. and Parker, L. E. (2013). Considering Inter-Task Resource Constraints in Task Allocation. *Journal of Autonomous Agents and Multi-Agent Systems*. Vol. 26.

Peer Reviewed Conference Publications

Zhang, Y., Sreedharan, S., Kulkarni, A., Chakraborti, T., Zhuo, H. and Kambhampati, S. (2016). Plan Explainability for Robot Task Planning. *Robotics: Science and Systems* (*RSS) Workshop on Planning for Human-Robot Interaction*.

Zhang, Y., Sreedharan, S., Kulkarni, A., Chakraborti, T., Zhuo, H. and Kambhampati, S. (2017). Plan Explainability and Predictability for Robot Task Planning. Under submission to *International Conference on Robotics and Automation (ICRA)*.

Zhang, Y., Sreedharan, S., Kambhampati, S. (2015). Capability Models and their applications in planning. *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS).*

Chakraborti, T., Zhang, Y., Kambhampati, S. (2016). Planning with Resource Conflicts in Human-Robot Cohabitation. *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS).*

Zhang, Y., Sreedharan, S., Kambhampati, S. (2016). A Formal Analysis of Required Cooperation in Multi-agent Planning. *International Conference on Automated Planning and Scheduling (ICAPS).*

Zhang, Y., Narayanan, V.,Chakraborti, T., and Kambhampati, S. (2015). A Human Factors Analysis of Proactive Support in Human-robot Teaming. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).*

Books & Book Chapters

Zhang, Y., Kim K., Fainekos, G. (2016). DisCoF: Cooperative Pathfinding in Distributed Systems with Limited Sensing and Communication Range. In Chong, Nak-Young and Cho, Young-Jo (Eds.), Springer Tracts in Advanced Robotics*, Distributed Autonomous Robotic Systems.* pp. 325-340.

Erin K. Chiou

Erin K. Chiou is an assistant professor of Human Systems Engineering at Arizona State University. She directs the Automation Design Advancing People and Technology (ADAPT) lab, which focuses on how automation can enhance cooperation to improve productivity, quality, and safety in our world. Dr. Chiou received her Ph.D. (2016) and M.S. (2013) in industrial and systems engineering from the University of Wisconsin-Madison, studying human factors and ergonomics, with a minor in health systems. She was an NSF Graduate Research Fellowship Program award recipient from 2013-2016 and a Graduate Engineering Research Scholars (UW-Madison) fellowship recipient. She has been a visiting researcher at the Wellness & Health Enhancement Engineering Lab at Northwestern University's Feinberg School of Medicine, and at New York University School of Medicine's Department of Population Health, with prior industry experience in healthcare with medical devices. In 2015, Dr. Chiou received a Best Article Award from the Human Factors and Ergonomics Society for her work on medication management with older adults. Her research interests include human-automation interaction with a focus on trust in automation, human-agent cooperation in complex systems, resilience engineering, and health systems. Her recent work involves using microworld environments to explore dyadic exchanges between humans and a synthetic teammate.

Select Publications

Chiou, E., Lee, J., Su, T. (submitted Oct 2016). Reciprocal exchange in human-agent cooperation. Targeted to *Ergonomics*.

Chiou, E., & Lee, J. (2016). Cooperation in human-agent systems to support resilience: A microworld experiment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. doi:10.1177/0018720816649094

Asan, O., Chiou, E., & Montague, E. (2015). Quantitative ethnographic study of physician workflow and interactions with electronic health record systems. *International Journal of Industrial Ergonomics*, *49*, 124–130. doi:10.1016/j.ergon.2014.04.004

Montague, E., & Chiou, E. (2014). Trust in complex work systems: A focus on information and communication technologies. In C. Korunka & P. Hoonakker (Eds.), *The Impact of ICT on Quality of Working Life* (pp. 143–152). Dordrecht: Springer.

Chiou, E., Venkatraman, V., Larson, K., Li, Y., Gibson, M., & Lee, J. (2014). Contextual design of a motivated medication management device. *Ergonomics in Design*, *22*(1), 8–15. doi:10.1017/CBO9781107415324.004

Montague, E., Xu, J., & Chiou, E. (2014). Shared experiences of technology and trust: An experimental study of physiological compliance between active and passive users in technology-mediated collaborative encounters. *IEEE Transactions on Human-Machine Systems*, *44*(5), 614–624. doi:10.1109/THMS.2014.2325859

Montague, E., Asan, O., & Chiou, E. (2013). Organizational and Technological Correlates of Nurses' Trust in a Smart Intravenous Pump. *Computers, Informatics, Nursing*, *31*(3), 142–149. doi:10.1097/NXN.0b013e3182812d95

Hajizadeh, N., Figueroa Perez, R., Uhler, L., Chiou, E., Perchonok, J., & Montague, E. (2013). Identifying design considerations for a shared decision aid for use at the point of outpatient clinical care: an ethnographic study at an inner city clinic. *Journal of Participatory Medicine*, *5*.

Mica R. Endsley

Mica Endsley is President of SA Technologies, a cognitive engineering firm and is the former Chief Scientist for the US Air Force. She has also held the position of Visiting Associate Professor at MIT in the Department of Aeronautics and Astronautics and Associate Professor of Industrial Engineering at Texas Tech University. She is Past-President of the Human Factors and Ergonomics Society. She received a Ph.D. in Industrial and Systems Engineering from the University of Southern California. Dr. Endsley is a recognized world leader in the design, development and evaluation of systems to support human situation awareness (SA) and decision-making, and the integration of humans and automation. She has authored over 200 scientific articles and is the co-author of "*Analysis and Measurement of Situation Awareness*" and "*Designing for Situation Awareness*".

Select Publications

Bolstad, C. A., & Endsley, M. R. (1999). Shared mental models and shared displays: An empirical evaluation of team performance. *Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA. pp. 213-217.

Bolstad, C. A., & Endsley, M. R. (2000). The effect of task load and shared displays on team situation awareness. *Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society,* Santa Monica, CA. pp. 189-192.

Bolstad, C. A., & Endsley, M. R. (2005). Choosing team collaboration tools: Lessons learned fron disaster recovery efforts. *Ergonomics in Design, Fall,* 7-13.

Endsley, M. R. (in press). From here to autonomy. *Human Factors*.

Endsley, M. R., & Jones, W. M. (2001). A model of inter- and intra- team situation awareness: Implications for design, training and measurement. In M. McNeese, E. Salas & M. Endsley (Eds.), *New trends in cooperative activities: Understanding system dynamics in complex environments* (pp. 46-67). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M. R., & Jones, D. G. (2012). *Designing for situation awareness: An approach to human-centered design* (2nd ed.). London: Taylor & Francis.

Endsley, M. R., Hansman, R. J., & Farley, T. C. (1998). Shared situation awareness in the flight deck - ATC system. *Proceedings of the AIAA/IEEE/SAE 17th Digital Avionics Systems Conference,* Bellevue, WA. pp.

Endsley, M. R., & Robertson, M. M. (2000). Situation awareness in aircraft maintenance teams. *International Journal of Industrial Ergonomics, 26*, 301-325.

Riley, J. M., & Endsley, M. R. (2004). The hunt for situation awareness: Human-robot interaction in search and rescue. *Paper presented at the Human Factors and Ergonomics Society 48th Annual Meeting,* Santa Monica, CA.

Riley, J. M., & Endsley, M. R. (2005). Situation awareness in HRI with collaborating remotely piloted vehicles. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, Santa Monica, CA. pp. 407-411.

Riley, J. M., Murphy, R. R., & Endsley, M. R. (2006). Situation awareness in the control of unmanned ground vehicles. In N. J. Cooke, H. L. Pringle, H. K. Pederson & O. Connor (Eds.), *Human factors of remotely operated vehicles* (pp. 359-372). Amsterdam: Elsevier.

Riley, J. M., Strater, L. D., Chappell, S. L., Connors, E. S., & Endsley, M. R. (2010). Situation awareness in human-robot interaction: Challenges and user interface requirements. In M. Barnes & F. Jentsch (Eds.), *Human-robot interaction in future military operations* (pp. 171-192). London: Ashgate.

14.0 Facilities

IRA A. FULTON SCHOOLS OF engineering
ARIZONA STATE UNIVERSITY

The Polytechnic School
Human Systems Engineering

Santa Catalina, Suite 150
7271 E. Sonoran Arroyo Mall
Mesa, AZ 85212

(480) 727-1338
poly.engineering.asu.edu

Nancy Cooke's Laboratory Facilities at ASU

New UAS Ground Station Simulator

Cooke's CERTT (Cognitive Engineering Research on Team Tasks) laboratory at ASU in Mesa, AZ consists of research simulators capable of providing context for six tasks: 1) Team-based Unmanned Aerial System (UAS) ground control, 2) Imagery analysis from a UAS, 3) Team-based mission planning and resource allocation, 4) Team-based cyber defense analysis, and 5) Human-Robot teaming, and 6)Team-based control of multiple heterogeneous unmanned underwater vehicles.

Simulators are based on abstractions of real world tasks using cognitive task analyses. Care is taken to preserve the fidelity of the cognitive and team aspects of the real tasks pertinent to the research questions central to the lab. The testbeds are well-equipped with metrics for assessing performance and cognitive state and provide for flexible scenario manipulation. Training and crew coordination research is typically conducted in the lab, but the equipment is also capable of testing workstation/interface configurations with humans in the loop.

Single Simulated Unmanned Aerial System Ground Control Workstation

Unmanned Underwater Vehicle Control Interface

Mission Planning Team Simulator

Additional Information:
Email: ncooke@asu.edu
https://webapp4.asu.edu/directory/person/559491
www.cerici.org
www.certt.com

Arizona State University
school of computing, informatics, & decision systems engineering

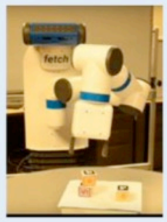
Brickyard Engineering (BYENG) 553
Address: 699 S. Mill Ave. Tempe, AZ 85281
Call: 480-965-3190 Email: cidse@asu.edu

Yochan Lab Facilities

Contact	Subbarao Kambhampati
Email	rao@asu.edu
Phone	+1 (480) 965 – 0113
Location	BYENG 560

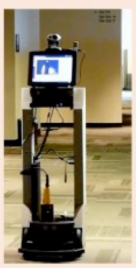
Rao's ***Yochan*** laboratory at ASU in Tempe, AZ provides state-of-the-art facilities for robotics and AI research. The group is particularly focused on challenges in human-aware AI and its applications to human-robot collaboration, and provides a unique perspective of algorithmic AI design, synthesis of natural robotic behaviors and human factors evaluation of the human-robot paradigm. Examples include explainable behavior in collaborative assembly tasks and coordination in search and rescue scenarios.

The robots in the group include Fetch, Baxter, PeopleBot, and several NAOs, each of which embody complementary capabilities useful across a wide range of human-robot interaction modalities; along with wearable technologies like Myo and Emotiv Epoc+ sensors that explore alternative dimensions of human-robot partnership.



The **Fetch** is a brand-new top-of-the-line mobile robot with a 7-DOF heavy duty arm. It provides precise perception, navigation and manipulation capabilities, and is useful in a variety of settings including collaborative assembly tasks, and search & delivery scenarios.

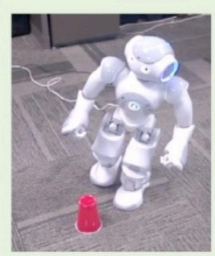
The **Baxter** is the most popular industry choice for collaborative assembly line production and assistance. The Baxter comes with two 7-DOF arms which are equipped with electric parallel grippers and vacuum cup grippers, and provide a gravity compensation mode that enables kinesthetic teaching through

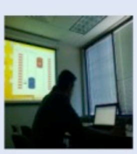


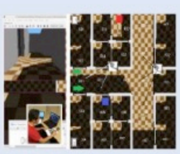
The **PeopleBot** is a mobile robot particularly built for interaction with humans via an easily accessible touch screen, camera and microphone on-board the Pioneer base. Built in a similar mold as the famous Cobots, the PeopleBot is the ideal choice as the in-house chaperone and delivery assistance.

Wearable technologies like the **Myo** sensor (EMG) and the **Emotiv Epoc+** headset (EEG) provide physiological clues to the response or attitude of humans towards their robotic teammates, as well as open up newer pathways of interaction.

The **NAO** robots are humanoid robots with access to bipedal motion, bimanual manipulation, speech and vision modules and special interaction modalities geared towards facilitating research in natural human-robot interaction. The NAOs have proved to be the #1 choice for human-robot interaction studies.







We also provide simulation test beds for human factor studies in settings such as autonomous driving and USAR tasks.

15.0 Special Test Equipment

None needed.

16.0 Equipment

No new equipment needed.

17.0 High Performance Computing Availability

Available but we won't be needing it for this project.

18.0 Budget & Justification

KEY PERSONNEL*

Nancy Cooke, Ph.D., **Principal Investigator,** (1.00 summer person months all years of the project). In this capacity she will oversee both the technical and administrative aspects of the project. She is responsible for the overall management of the project. Specifically, she will be responsible for coordinating all the studies, designing experimental protocols, training and supervising a postdoctoral scholar and two graduate students, overseeing analysis of results, and coordinating reporting and publication efforts. In addition, Dr. Cooke will organize regular team meetings for technical exchange and coordination of the overall effort.

Subbarao Kambhampati, Ph.D., Co-PI, (1.00 summer person months all years of the project). In this capacity he will assist with both overseeing both the technical and administrative aspects of the project, especially pertaining to the artificial intelligence thread. He will be responsible overseeing algorithm development, training and supervising two graduate students, and reporting results. In addition, Dr. Kambhampati will attend regular team meetings for technical exchange and coordination of the overall effort.

Erin Chiou, Ph.D., Co-Investigator, (1.00 summer person months all years of the project). Dr. Chiou will assist with the human-human and human robot studies, assisting with experimental design, development of the experimental protocol, data analysis and reporting. In addition, Dr. Chiou will attend regular team meetings for technical exchange and coordination of the overall effort.

Yu Zhang, Ph.D., Co-Investigator, (1.00 summer person months all years of the project). Dr. Zhang will assist with efforts relevant to the artificial intelligence thread. He will be responsible overseeing algorithm development, helping to train and supervise a graduate student, and reporting results. In addition, Dr. Zhang will attend regular team meetings for technical exchange and coordination of the overall effort.

Nathan McNeese, Ph.D., Senior Personnel, Post-Doctoral Associate (will allocate 2 months each year of the project**).** Dr. McNeese will assist with the human-human and human robot studies, assisting with experimental design, development of the experimental protocol, data analysis and reporting. In addition, Dr. McNeese will attend regular team meetings for technical exchange and coordination of the overall effort.

OTHER PERSONNEL*

(To Be Named), Graduate Student Research Assistant (GRA) (4). Funding is requested for the effort of four PhD students. For each year of the project, the GRA's will provide 100% of the allowable GRA effort (20 hours per week) for 9 academic months and 50% (Year's 1 & 2) of the allowable effort for 3 summer months. GRA no. 1 and no. 2 will be primarily focused on artificial intelligence work, helping to develop planning algorithms for robots based on empirical findings. GRA no. 3 and no. 4 will primarily focus on the Human-human and human-robot studies, assisting with design of the experiments and conducting data collection and analysis. All GRAs will assist with reporting and will attend regular meetings for technical exchange and coordination of the overall effort.

Fringe Benefits

Arizona State University defines fringe benefits as direct costs, estimates benefits as a standard percent of salary applied uniformly to all types of sponsored activities, and charges benefits to sponsors in accordance with the Federally-negotiated rates in effect at the time salaries are incurred. Benefit costs are expected to increase approximately 3% per year; the rates used in the proposal budget are based on the current Federally-negotiated Rate Agreement plus annual escalation for out years. $170,805 is the estimated cost of ERE for the personnel effort allocated in this project, which is based upon the following rates for FY 2018 and thereafter:

ERE Rate Estimates	Faculty	RA/TA	Post-doc
FY 2018 Estimated Rates	29.77%	14.73%	16.79%
FY 2019 Estimated Rates	30.66%	15.17%	17.29%
FY 2020 Estimated Rates	31.58%	15.63%	17.81%
FY 2021 Estimated Rates	32.53%	16.10%	18.34%
FY 2022 Estimated Rates	33.51%	16.58%	18.89%

TRAVEL

Domestic

Funding in the amount of $8,000 will be allocated for each project year to support domestic travel. The purpose of these trips is for the PI, senior personnel and students to attend annual AFOSR workshops to report on progress and scientific conferences to disseminate research results, engage with collaborators and the scientific community. These meetings will be held in various locations within in US.

Note: Using the applicable GSA or State Department per diem rates and current airfare rates and ASU authorized reimbursement rates (http://cfo.asu.edu/fs-travel-perdiem) the domestic travel budget includes 1-2 trips for 2-3 days for 1 traveler per year. Travel costs are budgeted according to ASU travel policy (FIN 500) located at http://www.asu.edu/aad/manuals/fin/fin509.html. ASU's travel system software provider, Concur Technologies, assesses a charge of $10.45/per person for each travel expense report submitted. The expense is a direct cost charged per trip.

Conference Illustration travel (for one person):

Domestic	3 Nights	4 days	1 traveler			
Boston, MA	Airfare	Hotel	Per Diem	Registration	Transportation & Misc.	Total
	$600	$356	$236	$500	$308	$2000
Domestic	3 Nights	4 days	1 traveler			
Washington DC	Airfare	Hotel	Per Diem	Registration	Transportation & Misc.	Total
	$650	$356	$236	$500	$258	$2000

<u>OTHER DIRECT COSTS</u>

Payments to Human Subjects: Funding ($2,000 per year) is requested to compensate human subjects (100-15 per year) for their participation in human-human and human-robot studies at the rate of $10 per hour for experimental sessions lasting between 2-3 hours.

Mica Endsley, Consultant ($10,000/yearly) will allocated towards her time to advise the project, especially on issues relevant to situation awareness, human-robot teaming, and USAF needs.

Graduate student research assistant (GRA) tuition support:
The total tuition remission costs associated with the graduate student(s) requested on this project are $404,448. Tuition remission is a mandatory benefit for graduate students and is charged in proportion to their expected effort on the project in accordance with ASU policy. The tuition charge for graduate students is $17,067 for FY18 and $18,099 for FY19, $20,929 for FY20 and $22,603 for FY21 and 24,411 for FY22. Tuition costs are expected to increase approximately 8% per year; hence, the rates used in the proposal budget are based on the current tuition rates escalated by 8% per year. Tuition charges are exempt from the Facilities and Administrative (F&A) costs.

Facilities and Administration (F&A, Indirect/Overhead) Costs:
ASU's indirect costs are calculated on Modified Total Direct Costs (MTDC) using F & A rates approved by the U.S. Department of Health and Human Services. The University's Rate Agreement specifies 54.5% for FY16 and provisional rates beyond July 1, 2016. The current rate agreement is dated May 13, 2016. MTDC comprises salaries and wages, fringe benefits, materials and supplies, services, travel, and subawards up to $25,000. Items excluded from F & A calculation include: graduate student tuition remission, participant support, subcontracts over the first $25,000, and capital equipment.

Budget Summary

Cooke-Budget Summary v5

Cost Categories	Period 1 1/1/17 12/31/18	Period 2 1/1/18 12/31/19	Period 3 1/1/19 12/31/20	Period 4 1/1/20 12/31/21	Period 5 1/1/21 12/31/22	Cumulative
Senior/Key Personnel:	**$72,066**	**$74,736**	**$93,031**	**$80,421**	**$83,450**	**$403,704**
Nancy Cooke	$14,033	$14,454	$14,888	$15,335	$15,795	$74,505
ERE:	$4,178	$4,432	$4,702	$4,988	$5,293	$23,593
Effort (FTE Months; AY/SUM/CAL):	0/1/1	0/1/1	0/1/1	0/1/1	0/1/1	
Subbarao Kambhampati	$20,389	$21,000	$21,630	$22,279	$22,948	$108,246
ERE:	$6,070	$6,439	$6,831	$7,247	$7,690	$34,277
Effort (FTE Months; AY/SUM/CAL):	0/1/1	0/1/1	0/1/1	0/1/1	0/1/1	
Erin Chiou	$10,000	$10,300	$10,609	$10,927	$11,255	$53,091
ERE:	$2,977	$3,158	$3,350	$3,555	$3,772	$16,812
Effort (FTE Months; AY/SUM/CAL):	0/1/1	0/1/1	0/1/1	0/1/1	0/1/1	
Yu Zhang	$11,111	$11,444	$23,576	$12,141	$12,506	$70,778
ERE:	$3,308	$3,509	$7,445	$3,949	$4,191	$22,402
Effort (FTE Months; AY/SUM/CAL):	0/1/1	0/1/1	1/1/02	0/1/1	0/1/1	
Other Personnel:	**$116,676**	**$120,639**	**$96,956**	**$100,277**	**$103,720**	**$538,268**
Post Doctoral Associate TBD02	$10,833	$11,158	$11,493	$11,838	$12,193	$57,515
ERE:	$1,819	$1,929	$2,047	$2,171	$2,303	$10,269
Effort (FTE Months; AY/SUM/CAL):	0/0/2	2/0/2	0/0/2	0/0/2	0/0/2	
Graduate Student TBD x 4 Ph.D's	$90,668	$93,384	$72,140	$74,304	$76,536	$407,032
ERE:	$13,356	$14,168	$11,276	$11,964	$12,688	$63,452
Effort (FTE Months; AY/SUM/CAL):	4.5/1.5/6	4.5/1.5/6	4.5/0/4.5	4.5/0/4.5	4.5/0/4.5	
Total Number Other Personnel	5	5	5	5	5	25
Total Salary, Wages and ERE:	**$188,742**	**$195,375**	**$189,987**	**$180,698**	**$187,170**	**$941,972**
Equipment:	**$0**	**$0**	**$0**	**$0**	**$0**	**$0**
Travel:	**$8,000**	**$8,000**	**$8,000**	**$8,000**	**$8,000**	**$40,000**
1. Domestic	$8,000	$8,000	$8,000	$8,000	$8,000	$40,000
2. Foreign	$0	$0	$0	$0	$0	$0
Participant/Trainee Support Costs:	**$0**	**$0**	**$0**	**$0**	**$0**	**$0**
1. Tuition/Fees/Health Insurance	$0	$0	$0	$0	$0	$0
2. Stipends	$0	$0	$0	$0	$0	$0
3. Travel	$0	$0	$0	$0	$0	$0
4. Subsistence	$0	$0	$0	$0	$0	$0
5. Other	$0	$0	$0	$0	$0	$0
6. Number of Participants/Trainees	0	0	0	0	0	0
Other Direct Costs:	**$83,768**	**$89,512**	**$90,600**	**$96,888**	**$103,680**	**$464,448**
1. Materials and Supplies	$0	$0	$0	$0	$0	$0
2. Publication Costs	$0	$0	$0	$0	$0	$0
3. Consulting Costs	$10,000	$10,000	$10,000	$10,000	$10,000	$50,000
4. ADP/Computer Services	$0	$0	$0	$0	$0	$0
5. Subaward/Consortium/Contractual	$0	$0	$0	$0	$0	$0
6. Equipment or Facility Rentals/User Fees	$0	$0	$0	$0	$0	$0
7. Alterations and Renovations	$0	$0	$0	$0	$0	$0
8. Tuition Remission	$71,768	$77,512	$78,600	$84,888	$91,680	$404,448
9. Human Subjects	$2,000	$2,000	$2,000	$2,000	$2,000	$10,000
10	$0	$0	$0	$0	$0	$0
Direct Costs:	**$280,510**	**$292,887**	**$288,587**	**$285,586**	**$298,850**	**$1,446,420**
Indirect Costs:	**$113,764**	**$117,379**	**$114,443**	**$109,380**	**$112,908**	**$567,874**
Total Direct and Indirect Costs:	**$394,274**	**$410,266**	**$403,030**	**$394,966**	**$411,758**	**$2,014,294**