# Shared Mental Models for Human-Robot Teams

**Julie A. Adams**
Department of Electrical Engineering
and Computer Science
Vanderbilt University
Nashville, TN USA
julie.a.adams@vanderbilt.edu

**Scott A. DeLoach**
Department of Computing
and Information Sciences
Kansas State University
Manhattan, KS USA
sdeloach@k-state.edu

**Matthias Scheutz**
Department of Computer Science
Tufts University
Medford, MA USA
matthias.scheutz@tufts.edu

**Abstract**

Shared mental models have been shown to improve human team performance. We thus conjecture that shared mental models (SMMs) integrated into cognitive robotic architectures might also improve the performance of mixed human-robot teams. To date, very little research has focused on developing appropriate computational constructs that can support domain independence and generalizability, while also being scalable. In this paper, we outline our proposed development of SMMs for cognitive robots.

## Introduction

Behavioral research has demonstrated that human teams coordinate activities more effectively and achieve better overall task performance when team members manage to track each other's goals, intentions, beliefs, as well as other mental and task-related states – a process called "shared mental models" (e.g., (Cannon-Bowers, Salas, and Converse 1993; Mathieu et al. 2000)). It is thus natural to assume that similar capabilities in robots, if properly integrated into the robots' computational architectures, might be able to improve the efficiency and productivity of human-robot teams.

*Shared mental models* (SMMs) (Cannon-Bowers, Salas, and Converse 1993) extend the concept of *mental models* (e.g., (Johnson-Laird 1983)) to teams based on the assumption that team members have various elements of their individual mental models in common. The early literature suggested four components of SMMs: *Equipment*, *Task*, *Team Interaction*, and *Team* models, which were consolidated into the the *Task* and *Team* models in the recent literature (e.g., (Mathieu et al. 2000)).

For robots to be able to use SMMs in the context of mixed human-robot teams, we need to both understand the representations and processes underwriting SMMs in humans and how those representations and processes can be captured as computational structures, i.e., what detailed information SMMs contain about team members, including their intentions, goals, knowledge, beliefs, capabilities, activities and performance. Unfortunately, past SMM research has almost exclusively focused on behavioral phenomena, such as the impact on team performance as well as the development and use of associated metrics (e.g., (Espevik, Johnsen, and Eid 2011; Mohammed, Ferzandi, and Hamilton 2010)). As a result, there are only a few suggestions on how to represent SMMs (e.g., using schemas (DuRussell and Derry 2005) or alternative representations (Lee, Johnson, and Jin 2012)).

Recent work in robotics has made some advances in the direction of SMMs. For example, a robot mental model based on cross training was demonstrated (Nikolaidis and Shah 2013), although the underlying assumptions limit scalability and do not support domain independence and generalizability. Others have suggested versions of robotic SMMs (Ososky et al. 2012) or developed simulated SMMs (Kennedy and Trafton 2007), but without providing a comprehensive framework for integrating SMMs into cognitive robotic architectures.

The following summarizes a proposed SMM formal framework that builds on the literature and how the framework can be integrated into cognitive robotic architectures.

## A Computational Framework

For robots to be able to use SMMs, we need to develop and integrate computational mechanisms into the cognitive robotic architecture that allow robots to monitor and predict human teammates' and adapt their behavior based on those predictions. For this purpose, we decompose SMMs into the *data representations* capturing information about the task, the other team members, and the environment, and the *computational processes* that operate on the data structures to create, maintain, revise and discard them. The former will likely differ from task to task, while the latter are intended to be more general mechanisms used across tasks.

We propose five comprehensive sets of predicates that capture different aspects of agents, tasks, and environments relevant to the SMM data representations: (R1) *agent capabilities and propensities*, including perceptions, actions, skills, traits, rank, and possibly other relevant agent properties; (R2) moment-to-moment *agent and task states*, including knowledge, belief and affective states, adopted goals and plans, and ongoing activities; (R3) known and accepted *obligations and norms* as they pertain to the task and performance domains, as well as general norms about agent behavior; (R4) *activity and equipment types*; and (R5) *functional roles of agents in teams* of the activities they can perform.

Many of the predicates have clear mappings to the SMM components. For example, the predicate *Achieve*$(\pi, \phi, \sigma)$ represents the achievement of a goal state $\phi$ via a plan $\pi$ in situation $\sigma$ maps directly to the SMM Task model, while the predicate *Perceivable*$(\alpha, X)$ (meaning that agent $\alpha$ can perceive $X$, where $X$ can be a type of agent, object, event, or activity) maps directly to the Team model.

We propose five corresponding sets of rules to update the data structures, while addressing how different predicates are related and how they can be used in the context of teams. For example, rules for negotiating goals and assignments with other team members are necessary, because goals can change or different goal assignments may become necessary during task execution due to unexpected events or the loss of capabilities by team members. This allows the formulation of various general principles for proposing plans, for example, that proposed plans to achieve common goals are adopted if no previous plans are in place. It is also necessary to be able to specify general principles of what is obligatory, what is permitted, and what is not permitted in the context of the team, task, and task environment. Overall, updates to a robot's SMM will be triggered by various events, mediated through the agent's perceptual system, and internal state changes. For example, $robot_A$ may perceive a new task-relevant object, which can trigger the instantiation of a new goal. $Robot_B$ observing $robot_A$ will also need to update its model to accommodate $robot_A$'s new goal.

## Integration Requirements

The integration of SMMs into cognitive robotic architectures must allow for efficient and consistent ways to capture, update and synchronize SMM state information, while using that information to drive robot behavior adaptation. Specifically, the integration needs to allow the robot to (1) directly observe humans to estimate their current performance fluctuations, (2) track human's progress toward their assigned goals, (3) determine the human's current task focus using performance estimates, goal assignments and progress, (4) estimate the human's cognitive workload based on task focus, (5) predict the implications of new/modified task assignments on the human using performance estimates and current task focus, and (6) adapt its behavior based on the measurements and predictions. Steps 1-4 represent *update* processes, while steps 5-6 represent *control* processes.

*Update* processes observe the environment and other teammates and use these observations to keep the SMM Team and Task state current. Such processes include those to update the equipment status, team and individual human task status, monitoring for new tasks and assignments, and updating environmental observations, such as noise, temperature, humidity, etc. Additionally, update tasks can be used to monitor human performance factors that can help predict human performance of assigned tasks. These update processes can be quite complex and make use of real-time observations as well as current performance estimates stored in the SMM.

*Control* processes use the SMM state information to control and adapt a robot's behavior and modify the team's organization. Using the current team and task state information, along with predictions of human performance, the robot can alter its individual behavior to adjust for human performance degradations or to augment a human during task executions. Robots can also use SMM state information to suggest task reassignments that support more effective team operations.

We have started integrating some of the above mechanisms into the DIARC architecture (Briggs and Scheutz 2012) and are developing a more sophisticated distributed version of SMMs to support asynchronous updates and distributed synchronization among multiple robots.

## References


Briggs, G., and Scheutz, M. 2012. Multi-modal belief updates in multi-robot human-robot dialogue interactions. In *Proceedings of AISB 2012*.

Cannon-Bowers, J. A.; Salas, E.; and Converse, S. 1993. Shared mental models in expert team decision making. In Castellan, N. J., ed., *Current issues in individual and team decision making*. Psychology Press. 221–246.

DuRussell, L. A., and Derry, S. J. 2005. Schema (mis)alignment in interdisciplenary teamwork. In Derry, S. J.; Schunn, C. D.; and Gernsbacher, M. A., eds., *Interdisciplinary Collaboration: An Emerging Cognitive Science*. Lawrence Earlbaum Associates, Publishers.

Espevik, R.; Johnsen, B. H.; and Eid, J. 2011. Outcomes of shared mental models of team members in cross training and high-intensity simulations. *Journal of Cognitive Engineering and Decision Making* 5(4):352–377.

Johnson-Laird, P. N. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consiciousness*. Cambridge: Cambridge University Press.

Kennedy, W. G., and Trafton, J. G. 2007. Using simulations to model shared mental models. In *Proceedings of the Eighth International Conference on Cognitive Modeling*, 253–254.

Lee, M.; Johnson, T.; and Jin, M. H. 2012. Toward understanding the dynamic relationship between team and task shared mental models as determinants of team and individual performances. *International Journal of Information Technology and Business Management* 8(1).

Mathieu, J. E.; Goodwin, G. F.; Heffner, T. S.; Salas, E.; and Cannon-Bowers, J. A. 2000. The influence of shared mental models on team process and performance. *Journal of Applied Psychology* 85(2):273–283.

Mohammed, S.; Ferzandi, L.; and Hamilton, K. 2010. Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management* 36(4):876–910.

Nikolaidis, S., and Shah, J. A. 2013. Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, 33–40.

Ososky, S.; Schuster, D.; Jentsch, F.; Fiore, S.; Shumaker, R.; Lebiere, C.; Kurup, U.; Oh, J.; and Stentz, A. 2012. The importance of shared mental models and shared situation awareness for transforming robots form tools to teammates. In *Proceedings of SPIE 8387, Unmanned Systems Technology XIV*, 838710–1.