5-4-2022

Factored Beliefs for Machine Agents in Decentralized Partially Observable Markov Decision Processes

Joshua Lapso

Gilbert L. Peterson
Air Force Institute of Technology

Factored Beliefs for Machine Agents in Decentralized Partially Observable Markov Decision Processes *

Joshua A. Lapso and **Gilbert L. Peterson**
Air Force Institute of Technology
2950 Hobson Way
Wright-Patterson Air Force Base, OH 45433

Abstract

A shared mental model (SMM) is a foundational structure in high performing, *task-oriented teams* and aid humans in determining their teammate's goals and intentions. Higher levels of mental alignment between teammates can reduce the direct dialogue required for team success. For *decision-making teams*, a transactive memory system (TMS) offers team members a map of specialized knowledge, indicating source of knowledge and the source's credibility. SMM and TMS formulations aid human-agent team performance in their intended team types. However, neither improve team performance with a project team–one that requires both behavioral and knowledge integration. We present a hybrid cognitive model (HCM) for machine agents that subsumes the integrated portions of a team's transactive memory in an SMM. The unified structure of the HCM enables contextual switches during execution for machine agents, over the two cognitive formulations with comparable computational complexity of a single cognitive model. Results in a multi-agent project environment demonstrates how the HCM provides machine agents with a generalizable cognitive structure that is able to maintain fully factored belief states with minimal inter-agent communication.

Studies in the human factors and cognitive psychology disciplines provide details with regard to high performing teams. Effective teams transcend complex domain uncertainty by achieving an emergent state of shared cognition, in which knowledge is organized, represented, and distributed to team members for rapid execution (DeChurch and Mesmer-Magnus 2010). Much of the work in modeling shared cognition leverages either the concept of a share mental model (SMM) (Cannon-Bowers, Salas, and Converse 1993) or a transactive memory system (TMS) (Wegner, Giuliano, and Hertel 1985; Moreland, Argote, and Krishnan 1996). SMMs are joint understandings of tasks (e.g., temporal sequencing of actions, arrangement of materials, or resource interdependencies) shared among team members. Alternatively, a TMS offers team members a global context of informa-

tion, mapping where knowledge resides and credibility of the knowledge holder.

Team types are usually categorized as either action teams, decision-making teams, or project teams (DeChurch and Mesmer-Magnus 2010). Action teams coordinate and perform physical tasks such as a sports team or first responders. Decision-making teams (e.g., company boards, medical diagnostic panels, or financial planners) process information and provide expert opinions. Project teams incorporate both action and decision-making teams. In the project team setting, human agents perform contextual switches between processing tasks and processing information. Project teams include research and development centers, military task forces, and disaster recovery teams. This work focuses on the mental models of machine agents in project teams.

This work investigates foundational characteristics of the HCM that incorporates aspects of both SMMs and TMSs, enabling emergent states (i.e., shared cognition) for machine agents. The HCM forms an SMM by subsuming only the integrated portions of the TMS and treating the differentiated portions as auxiliary knowledge. The multi-agent decision problem (MADP) framework (Oliehoek and Amato 2016) enables generalizations elusive in more rigid models. The HCM's embedded cognitive elements augment internal state updates with intent projections—inferences about team member intentions formulated by attending over elements in the TMS and SMM.

Evaluation of the HCM is conducted through direct comparison with a decentralized partially observable Markov decision process (Dec-POMDP)[1] (Bernstein et al. 2002). The performance of the HCM and Dec-POMDP are first evaluated in terms of average cumulative reward over concurrent trials. Subsequently, the accuracy of machine agents with factored belief states in the HCM are assessed against the joint beliefs of Dec-POMDP agents.

Related Work

Investigations into the procedural, psychological, and strategic factors of high performing teams are conducted from numerous perspectives. Those relevant to the formulation

---

[1] Equivalently referenced as a multi-agent POMDP (MPOMDP) (Oliehoek and Amato 2016) or multi-agent team decision problem (MTDP) (Pynadath and Tambe 2002)

of the HCM are shared cognition (Wegner 1987; Cannon-Bowers, Salas, and Converse 1993) and multi-agent systems (MAS) (Wooldridge 2009; Weiss 2013; Shoham and Leyton-Brown 2008). This sections presents a brief introduction to the foundational perspectives used to formulate the HCM for machine agents.

Shared Cognition

A phenomena in high performing teams is how individuals transform into more than the sum of individual skill and knowledge (Corning 2002). This transformation is typically detailed by an integration of individual minds which form a shared cognitive state. Two prominent theories of cognition and group mind are the SMM (Cannon-Bowers, Salas, and Converse 1993) and TMS (Wegner, Giuliano, and Hertel 1985) respectively. These theories are differentiated by their principal characteristics[2] (see Table 1) overlapping only on the concept of individual and team member intent. Both theories have shown through independent studies (DeChurch and Mesmer-Magnus 2010; Ren and Argote 2011) critical contributions toward an understanding of the emergent state of shared cognition between individual team members.

Table 1: Characteristics[2] of SMMs and TMSs.

Characteristic[2]	Description	Model
Task Knowledge	Known completion requirements (T_R), skills (T_S), sequence (T_Q), and constraints (T_C) required for success	SMM[ab]
Goals (G)	A shared understanding of what individuals, teams, or coalition aim to achieve	SMM[bc]
Roles (R)	How the team is led, organized, and currently tasked	SMM[bc]
State (S)	Current understanding of the environment, available resources, progress towards goals, and limitations	SMM[bc]
Shared Communication (C)	A collection of methods for message passing between members, including how and when to employ each	SMM[bc]
Plans (P)	Proposals for action established for anticipated events and dealing with uncertainty	SMM[bc]
Intentions (I)	Next action, decision, or focused attention of each team member	Both[bcde]
Knowledge Map	Direct elements of information (T_R, T_S, T_Q, T_C, G, R, S, P, and I) or a location (L) where elements reside (or do not) within the team	TMS[cde]
Credibility Map	A knowledge map overlay assigning credibility to a knowledge holder	TMS[cde]
Knowledge Organization	A schema for how information is stored	TMS[cde]
Knowledge Accessibility	An understanding of who can acquire and the process for acquiring information internal/external to the team	TMS[cde]

SMMs are explanatory mechanisms that unify team understanding and ultimately improve team performance (Cannon-Bowers and Salas 2001; DeChurch and Mesmer-Magnus 2010). Cannon-Bowers, et al. (1993) showed mental model alignment reduces direct dialogue and increases communication efficiency in several studies (Cannon-Bowers and Salas 2001). SMM alignment does not consider levels of expertise. For SMMs, alignment is a binary factor assessing whether or not a member posses the desired task-focused or team-focused knowledge (DeChurch and Mesmer-Magnus 2010). Task-focused knowledge includes skills and equipment employed by a team (e.g., Task Skill (T_S) or State (S) from Table 1). Team-focused knowledge is concerned with team characteristics (e.g., Goals (G), Shared Communications (C), Plans (P), and Intentions (I) from Table 1). Fully aligned SMMs are similar in form to fully integrated TMSs, where each member posses all relevant knowledge independently. However, TMSs leverage the intimate relationship context of a team.

Transactive memory theorists propose that memories are first encoded in individual memory space and then re-encoded into transactive memory space through interactions, but only in the *context* of a unique small group of individuals. The initial study of transactive memory (Wegner, Giuliano, and Hertel 1985) focused on human dyads in intimate relationships, but later studies extended the theory to small groups (Wegner 1987) and teams (Moreland, Argote, and Krishnan 1996). Consider a team context where one member with expert knowledge or in an established role may have experienced the same information differently than other members. Through interaction and dialogue, team members re-encode individual memories into a TMS, often generating new, more complete representations, in the joint understanding of all individual perspectives. Knowledge in a TMS can be differentiated or integrated (Wegner, Giuliano, and Hertel 1985) both have beneficial contributions to shared cognition, but Wegner (1987) regards both as important to teamwork. A purely differentiated TMS would create information bottlenecks or in the worst case, a single point of failure for the team. On the other hand, a fully integrated TMS required overhead proportional to team size and could foster indecision (Wegner, Giuliano, and Hertel 1985). Striking a balance between integration and differentiation ensures a prudent amount of knowledge duplication among members, enabling simultaneous tasks work and multilateral decisions.

Multi-agent Systems

SMMs are ubiquitous throughout MAS implementations (Wooldridge 2009; Weiss 2013; Shoham and Leyton-Brown 2008) as Fan and Yen (2004) detail. While the TMS literature spans several scientific disciplines (Ren and Argote 2011), there remains a dearth of implementations for machines agents in a MAS (Corbett 2012). Existing MAS algorithmic model types often incorporate communication and cognitive elements as auxiliary observations in fully defined multi-agent MDP variations (e.g., Dec-MDP, MPOMDP, and Dec-POMDP-COM) (Oliehoek and Amato 2016). Multi-agent models that construct and maintain

---

[2] The term characteristic is used to describe elements of either model critical to functionality. The TMS knowledge map may subsume items in an SMM, but the element types themselves are not critical to knowledge map functionality and thus not considered principal characteristics of TMSs.

---

[a](Cannon-Bowers, Salas, and Converse 1993)
[b](Fernandez et al. 2017)
[c](DeChurch and Mesmer-Magnus 2010)
[d](Wegner 1987)
[e](Moreland, Argote, and Krishnan 1996)

full joint action-observation histories, become intractable for problems with large observation spaces and infinite horizons (Oliehoek and Amato 2016; Murphy 2002; Boutilier 1996). It is also understood that Dec-POMDPs preclude agents from accessing a Markovian signal during operation and in turn preventing them from forming individual beliefs.

The Markov Multi-Agent Environment (MME) (Oliehoek and Amato 2016) enables a solution analogous to the human-agent cognitive models. The MME is a generalized, partially defined Markov model that provides researchers flexibility in completing the MAS model definition with dynamically defined agent models. By decoupling machine agents from the environment, the MME reaps the benefits of Dynamic Bayesian Network (DBN) representations. The HCM defines machine agent models as fully factored DBNs (Murphy 2002) with an additional visible layer for the TMS, and by subsumption the SMM, between the hidden and belief states (Åström 1965). Within the HCM optimal action selection and intent projection policies for machine agent teams are constructed with temporal difference learning (Sutton and Barto 2018) algorithms.

Methodology

This research proposes that for a machine agent within a task-oriented team, the SMMs act as a filter for team-relevant items within a TMS. Common knowledge items (e.g., Task Skills (T_S), State (S), Goals (G), Shared Communications (C), Plans (P), and Intentions (I) from Table 1) are collected in the intersections of transactive memory as shown in Figure 1. Essentially, knowledge in the TMS is known by one (*fully differentiated*), some (*partially integrated*), or all (*fully integrated*) of the team members. When the team is presented with task work, the fully integrated portion of the TMS is an SMM. Alternatively, if the team is presented knowledge work (i.e., an expert opinion or decision), the differentiated and integrated portions are leveraged.

Consider a three member team as illustrated in Figure 1. The basis for this cognitive model is a TMS with both integrated and differentiated portions. Each member possesses knowledge and credibility maps as well as information on how offline or external knowledge is organized and accessed. These elements are trainable parameters and are updated as new knowledge is learned or team members change. For instance, the credibility rating of a member could decrease if that member's input results in a poor decision. Likewise, a member's credibility rating could increase as they are observed to successfully complete tasks. An SMM is formed by elements (T_R, T_S, T_Q, T_C, G, R, S, C, P, and I) in the integrated and partially integrated portion of the TMS, which can be thought of as an intersection of the team's transactive memory. The differentiated portion of the TMS is not used during task work, but is fully utilized in decision making. Artificial agents may transfer learned parameters as individual memory when instantiated within other teams, analogous to a human agent being assigned to another team.
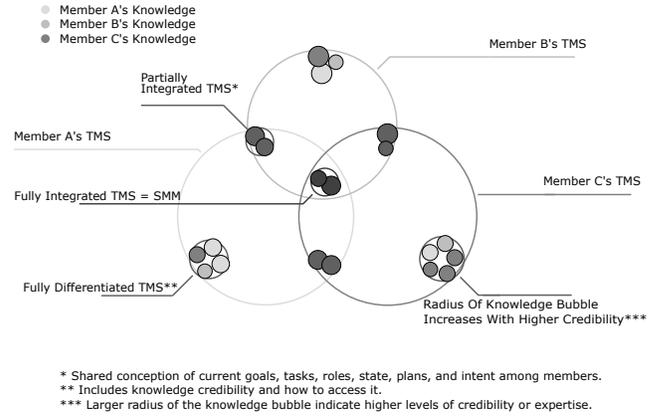

Figure 1: The SMM is formed in the intersections of the TMS. Task skill knowledge (T_S), Goals (G), Plans (P), and Intentions (I) are may be common to some or all team members, especially when training occurred prior to teamwork execution. Knowledge map location information (L) is present in transactive memory, but not contained in the intersection and thus not a part of the SMM.
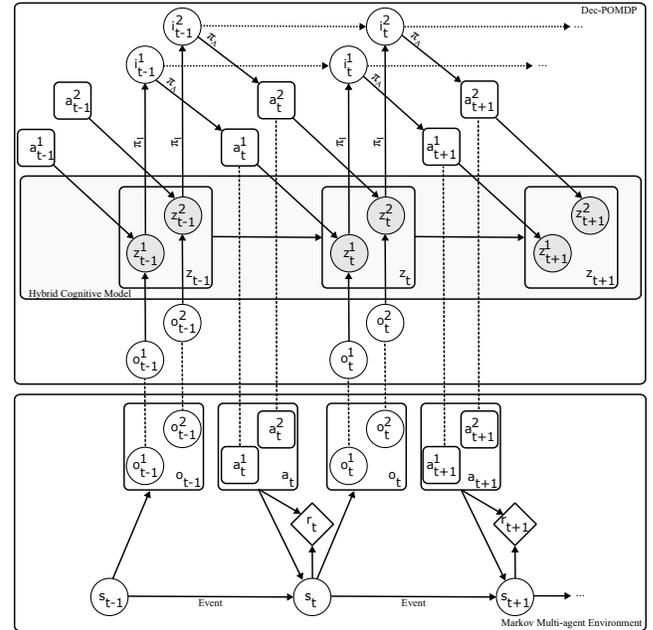

Figure 2: The HCM as a DBN for multiple agents. The joint environmental state transitions are indicated by the sequence s_{t-1}, s_t, s_{t+1}, ... Agent one's observations are the sequence o_{t-1}^1, o_t^1, ..., and actions a_{t-1}^1, a_t^1, ..., while agent two's observations are indicated by the sequence o_{t-1}^2, o_t^2, ..., and actions by the sequence a_{t-1}^2, a_t^2, ... Agent TMSs are denoted z^1 and z^2 for agents one and two respectively. The dotted edges denote internal state transition, while the dashed lines between observation and action nodes are used to indicate that observations are replicated into the agent model and actions from the agent model into the MME.

Model Formulation

The hybrid cognitive model is a joint, event-driven state space represented as a DBN, shown for the two agent case in Figure 2. Each agent maintains a fully factored TMS (z^n),

which is integrated through agent action and communication and subsumed in an SMM. Eight types of information are specified for this DBN:

1. The prior internal state distribution over the state variables, $I(t = 0) = P(i_0)$

2. The state transition model, $T(s_t, s_{t-1}, a_t) = P(s_t \mid s_{t-1}, a_t)$

3. The observation model, $O(o_t, s_t) = P(o_t \mid s_t)$

4. The internal state model, $I(i_t, i_{t-1}, o_t) = P(i_t \mid i_{t-1}, o_t)$

5. The reward function, $R(s_t, a_{t+1})$

6. The stochastic action selection policy $\pi_A : I \to \Delta(A_i)$

7. The stochastic intent projection policy $\pi_I : I \to \Delta(Z_i)$

8. The off-policy learning function, $Q(s_t, a_t, z_{t-1}) = Q(s_t, a_t, z_{t-1}) + \alpha[R(s_t, a_t) + \gamma \max_{a' \in \mathcal{A}} Q(s_{t+1}, a', z_t) - Q(s_t, a_t, z_{t-1})]$

For smaller state spaces the MME, utilizing the HCM's agent model, is solvable directly as detailed in (Oliehoek and Amato 2016)—requiring only minor algorithmic changes to incorporate the TMS. However, as Dec-POMDP complexity is exponential in both the number of agents and length of horizon, a direct solution becomes intractable for realistic state spaces. This work instead employs Bayesian filtering techniques to approximate internal state. Algorithm 1 presents the generalized form of the HCM.

Algorithm 1 Direct inference with discrete Bayesian filter

Function HYBRIDCOGNITIVEMODEL():
Variables:
    $\{M\}$: A set of m machine agent
    $\{z_0\}$: The TMS containing a priori knowledge for m agents
    s_0: The initial state of the domain
    $\{i_0\}$: The set of prior belief distribution over s_t
    π_A: The action selection policy
    π_I: The intent projection policy
    $\overline{i_t^m}$: The predicted internal state distribution
    i_t^m: The corrected internal state distribution
$t \leftarrow 0$
$s_{t\text{-}1} \leftarrow s_0$
$i_{t\text{-}1}^m \leftarrow i_0$
$z_{t\text{-}1}^m \leftarrow z_0$
repeat
    $t \leftarrow t + 1$
    $i_{max}^m \leftarrow \underset{t\text{-}1}{\operatorname{argmax}}(i_{t\text{-}1}^m))$
    Select a_t^m from $\pi(i_{max}^m\ a_t^m)$
    Transition to s_t, and observe o_t^m
    $\overline{i_t^m} \leftarrow \sum_{t\text{-}1} P(i_t^m \mid a_t^m, i_{t\text{-}1}^m) P(i_{t\text{-}1}^m)$
    $i_t^m \leftarrow \alpha P(o_t^m \mid i_t^m) P(\overline{i_t^m})$
    $\pi_A \leftarrow Q(s_t, a_t, z_{t-1})$
    $\pi_I \leftarrow Q(s_t, a_t, z_{t-1})$
until Execution phase is terminated

Evaluation

The first experiment evaluates the performance of two models with average cumulative reward as the optimality criterion. Fully factored HCM agents and a Dec-POMDP are instantiated over identical MMEs for an equivalent number of training episodes. Three team configurations are considered in both models; (1) homogeneous agents with no specialization, (2) heterogeneous agents where a single member has
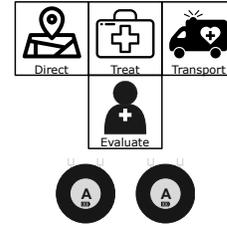

Figure 3: A two agent visualization of the Triage problem domain

specialization, and (3) homogeneous agents with specialization. After training the performance of both models are recorded and analyzed.

The second experiment considers the accuracy of the factored belief states of HCM with regard to the joint belief states of a Dec-POMDP. The HCM agents are instantiated alongside a collective entity that has visibility over joint actions and observations (i.e., a Dec-POMDP). Belief state histories for both agents and the agent collective are extracted for analysis. This experiment also utilizes the three team configurations specified above.

The Triage Domain

The Triage problem draws inspiration from the Decentralized Tiger problem (Nair et al. 2003). Two or more machine agents must triage victims of a natural disaster and determine the best method for getting them safely to the Rendezvous location. The state variable, $s_t \in \{Fair,\ Serious,\ Critical,\ Terminal\}$ represents the victims condition. Machine agents have four actions $a_t \in \{Evaluate,\ Direct,\ Treat,\ Transport\}$ and four observations ($o_t \in \{Minor,\ Moderate,\ Serious,\ Critical\}$) over the severity of injuries.

Agents in this domain are presented with victims and must make repetitive life or death decisions as illustrated in Figure 3. Victims in fair condition can safely walk to the Rendezvous location, but those in serious conditions may incur further injuries. The walk is fatal to victims in critical condition. Treating a victim in fair condition has no effect on their state and wastes resources. A victim in serious condition may have their state transition to minor with treatment. Treatment may improve the state of a victim in critical condition, but it may also be fatal. Transporting victims in fair or serious condition makes the resource unavailable for untriaged victims in critical condition. Transporting Victims in critical condition increases their probability of survival. When victims are evaluated jointly, the knowledge and credibility maps for all agents involved are leveraged in the decision. However, divergent actions may cause confusion in the victim leading to undesirable behavior. Repeated evaluations may improve accuracy of diagnosis, but the delay may be fatal to victims in critical condition.

The Triage domain's state transition model is captured in Table 2. The direct and transport actions result in a terminal state; either the victim perishes or makes a full recovery. Other actions can upgrade or downgrade a victims condition. Victims in critical conditions may perish during treatment or

Table 2: Triage domain state transition probabilities.

Joint Action	State Transitions									
	$S_{Crit} \to S_{Term}$	$S_{Crit} \to S_{Crit}$	$S_{Crit} \to S_{Ser}$	$S_{Ser} \to S_{Crit}$	$S_{Ser} \to S_{Ser}$	$S_{Ser} \to S_{Fair}$	$S_{Ser} \to S_{Term}$	$S_{Fair} \to S_{Fair}$	$S_{Fair} \to S_{Term}$	All Others
⟨evaluate − evaluate⟩	0.25	0.75	0.00	0.10	0.90	0.00	0.00	1.00	0.00	0.00
⟨evaluate − treat⟩	0.25	0.25	0.50	0.00	0.50	0.50	0.00	1.00	0.00	0.00
⟨treat − evaluate⟩	0.25	0.25	0.50	0.00	0.50	0.50	0.00	1.00	0.00	0.00
⟨treat − treat⟩	0.05	0.25	0.70	0.00	0.30	0.70	0.00	1.00	0.00	0.00
otherwise	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00

Table 3: Triage domain reward function.

Joint Action	Victim State		
	S_{Crit}	S_{Ser}	S_{Fair}
⟨evaluate − evaluate⟩	-2	-2	-2
⟨evaluate − direct⟩	-101	14	24
⟨evaluate − treat⟩	4	-6	-11
⟨evaluate − transport⟩	24	9	-26
⟨direct − evaluate⟩	-101	14	24
⟨direct − direct⟩	-100	30	50
⟨direct − treat⟩	-95	10	15
⟨direct − transport⟩	-75	25	0
⟨treat − evaluate⟩	4	-6	-11
⟨treat − direct⟩	-95	10	15
⟨treat − treat⟩	10	-10	-20
⟨treat − transport⟩	30	5	-35
⟨transport − evaluate⟩	24	9	-26
⟨transport − direct⟩	-75	25	0
⟨transport − treat⟩	30	5	-35
⟨transport − transport⟩	50	20	-50

Table 4: Observation probabilities for trained roles formatted as $[Critical, Serious, Moderate, Minor]$. There are no observations over Terminal states since the domain resets.

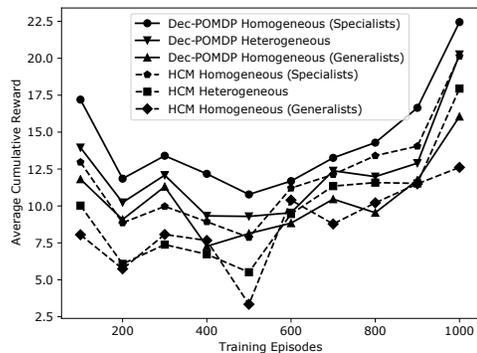| Victim Condition	Assigned Role	
	Diagnostic Specialist	Generalist
Critical	[0.90, 0.07, 0.02, 0.01]	[0.55, 0.30, 0.10, 0.50]
Serious	[0.09, 0.80, 0.09, 0.02]	[0.10, 0.40, 0.40, 0.10]
Fair	[0.01, 0.04, 0.35, 0.60]	[0.05, 0.10, 0.03, 0.55]

Figure 4: Average cumulative rewards for Dec-POMDP and HCM agent configurations based on number of training episodes.

evaluation. Finally, treatments and evaluations have no effect on victims in fair condition.

All rewards in this domain are shared and cumulative. A positive reward is generated when a victim is successfully triaged as shown in Table 3. A comparable penalty is incurred for actions that put the victims at risk.

With no observations available at time t_0 the initial agent actions a_0^1 and a_0^2 are conditioned upon the elements of their HCM established a priori, z_0^1 and z_0^2, respectively. For the initial instantiation of the HCM, the knowledge elements within TMSs are limited to $T_S \in \{diagnostics, \ treatments\}$, $T_C \in \{transportAvailible, \ transportUnavailible\}$, and $R \in \{diagnosticSpecialist, \ generalist\}$.

The observation probabilities in Table 4 represent the likelihood of agents in a trained role observing an injury level, based on a victims actual state.

Results

The Dec-POMDP team configurations outperformed the HCM in pair-wise comparison of their average cumulative reward as illustrated in Figure 4. However, this is expected as

the Dec-POMDP models assume full joint knowledge over the action-observation histories. The HCM configurations with specialized homogeneous agents and heterogeneous agents performed similarly to heterogeneous and generalized homogeneous Dec-POMDP agents, respectively. The generalized homogeneous HCM agents displayed high levels of variability and the agent's policies did not converge within 1000 training episodes.

Team configurations were then evaluated on their accuracy of belief state tracking. Figure 5 shows the Dec-POMDP team configurations were more accurate through pair-wise comparison of team configurations, but the HCM teams remained within a reasonable margin–often less than ten percent. The overall performance by the HCM teams demonstrates the efficacy of fully factored belief states on cooperative MAS domains.
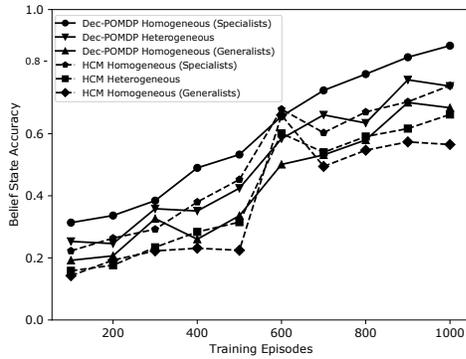
Figure 5: Belief state tracking accuracy for Dec_POMDP and HCM agent configurations based on number of training episodes.

Conclusion and Future Work

This work introduced a cognitive model, leveraging characteristics of both SMMs and TMSs, allowing cooperative machine agents to form and maintain fully factored belief states. Three HCM configurations were evaluated alongside an equivalent Dec-POMDPs by comparing average cumulative reward and accuracy of belief state tracking. The analysis concluded that the HCM enables machine agents to operate effectively in a purely decentralized capacity.

The next phase of research should test the HCM in broader environments that incorporates knowledge work as well as adversarial challenges. In the broader state spaces direct Bayesian filtering is not tractable, but internal state approximations with particle filters (Thrun, Burgard, and Fox 2005; Russell, Norvig, and Davis 2010) are a promising alternative. Future research should also include an evaluation of the HCM's efficacy in human-machine teams for both cooperative and adversarial domains.

References

Åström, K. J. 1965. Optimal control of Markov Processes with incomplete state information. *Journal of Mathematical Analysis and Applications* 10:174–205.

Bernstein, D. S.; Givan, R.; Immerman, N.; and Zilberstein, S. 2002. The complexity of decentralized control of markov decision processes. *Mathematics of Operations Research* 27(4):819–840.

Boutilier, C. 1996. Planning, learning and coordination in multiagent decision processes. In *Proceedings of the 6th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK '96, 195–210. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Cannon-Bowers, J. A., and Salas, E. 2001. Reflections on shared cognition. *Journal of Organizational Behavior* 22(2):195–202.

Cannon-Bowers, J.; Salas, E.; and Converse, S. 1993. Shared mental models in expert team decision making individual and group decision making. *Lawrence Erlbaurm* 221–246.

Corbett, A. 2012. *Agent-based Modelling of Transactive Memory Systems and Knowledge Processes in Agile versus Tradi-*

tional Software Development Teams. Ph.D. Dissertation, The University of Sheffield.

Corning, P. 2002. The re-emergence of "emergence": A venerable concept in search of a theory. *Complexity* 7:18–30.

DeChurch, L., and Mesmer-Magnus, J. 2010. Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research, and Practice* 14:1–14.

Fan, X., and Yen, J. 2004. Modeling and simulating human teamwork behaviors using intelligent agents. *Physics of life reviews* 1(3):173–201.

Fernandez, R.; Shah, S.; Rosenman, E. D.; Kozlowski, S. W.; Parker, S. H.; and Grand, J. A. 2017. Developing team cognition: a role for simulation. *Simulation in healthcare: journal of the Society for Simulation in Healthcare* 12(2):96.

Moreland, R.; Argote, L.; and Krishnan, R. 1996. Socially shared cognition at work: Transactive memory and group performance.

Murphy, K. P. 2002. *Dynamic Bayesian Networks: Representation, Inference and Learning*. Ph.D. Dissertation, USA.

Nair, R.; Tambe, M.; Yokoo, M.; Pynadath, D.; and Marsella, S. 2003. Taming decentralized POMDPs: Towards efficient policy computation for multiagent settings. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, IJCAI'03, 705–711. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Oliehoek, F. A., and Amato, C. 2016. *A Concise Introduction to Decentralized POMDPs*. Springer Publishing Company, Incorporated, 1st edition.

Pynadath, D. V., and Tambe, M. 2002. The communicative multiagent team decision problem: Analyzing teamwork theories and models. 16(1):389–423.

Ren, Y., and Argote, L. 2011. Transactive memory systems 1985–2010: An integrative framework of key dimensions, antecedents, and consequences. *The Academy of Management Annals* 5(1):189–229.

Russell, S.; Norvig, P.; and Davis, E. 2010. *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence. Prentice Hall.

Shoham, Y., and Leyton-Brown, K. 2008. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press.

Sutton, R. S., and Barto, A. G. 2018. *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book.

Thrun, S.; Burgard, W.; and Fox, D. 2005. *Probabilistic Robotics*. Intelligent Robotics and Autonomous Agents series. MIT Press.

Wegner, D. M.; Giuliano, T.; and Hertel, P. T. 1985. Cognitive interdependence in close relationships. In *Compatible and incompatible relationships*. Springer. 253–276.

Wegner, D. M. 1987. *Transactive Memory: A Contemporary Analysis of the Group Mind*. New York, NY: Springer New York. 185–208.

Weiss, G. 2013. *Multiagent Systems*. The MIT Press, 2nd edition.

Wooldridge, M. 2009. *An Introduction to MultiAgent Systems*. 2nd edition.