Learning complementary action with differences in goal knowledge

Jeremy Karnowski (jkarnows@cogsci.ucsd.edu)
Department of Cognitive Science, 9500 Gilman Drive
La Jolla, CA 92093-0515 USA

Edwin Hutchins (ehutchins@cogsci.ucsd.edu)
Department of Cognitive Science, 9500 Gilman Drive
La Jolla, CA 92093-0515 USA

Abstract

Humans, as a cooperative species, need to coordinate in order to achieve goals that are beyond the ability of one individual. Modeling the emergence of coordination can provide ways to understand how successful joint action is established. In this paper, we investigate the problem of two agents coordinating to move an object to one agent's target location through complementary action. We formalize the problem using a decision-theoretic framework called Decentralized Partially Observable Markov Decision Processes (Dec-POMDPs). We utilize multi-agent Q-learning as a heuristic to obtain reasonable solutions to our problem and investigate how different agent architectures, which represent hypotheses about agent abilities and internal representations, affect the convergence of the learning process. Our results show, in this problem, that agents using external signals or internal representations will not only eventually perform better than those that are coordinating in physical space alone but also outperform agents that have independent knowledge of the goal. We then employ information theoretic measures to quantify the restructuring of information flow over the learning process. We find that the external environment state varies in its informativeness about agents' actions depending on the agents' architecture. Finally, we discuss how these results, and the modeling technique in general, can address questions regarding the origins of communication.

Keywords: Dec-POMDPs; multi-agent Q-learning; Behavioral Info-Dynamics; mutual information

Introduction

The moment we move from a study of individual cognition to a detailed analysis of the social realm, we have committed ourselves to the investigation of a different type of system. There is no centralized controller; this system is inherently decentralized. The questions we ask, however, may be similar. Just as we wish to study how a individual decision maker adapts its behavior in a task environment, we can investigate the ways in which multiple, possibly non-identical, decision makers reorganize their internal world and their external interactions to form a new functional system that solves a problem which cannot be addressed by one individual alone (Hutchins, 1995).

One important problem that cooperative agents face is how to coordinate their movements to arrive at a goal known only to one of the agents. This problem was addressed in Hazlehurst and Hutchins (1998), where the authors constructed an algorithm that allowed for a set of agents to converge on similar form-meaning mappings which also related to their movements within a given environment. This setup, like many modeling studies that focus on issues of hidden goals of other agents, has a strong predilection towards imitative learning. Not all learning and reorganization in a multi-agent system is imitative, however, and another focus of modeling should be on complementary action learning (Hutchins & Johnson, 2009). It has been shown elsewhere that agents can learn to coordinate in complementary ways without sharing information about each other (Sen, Sekaran, Hale, et al., 1994), but this presumes an environment where there is only one destination and both agents know its identity. By combining aspects from these two studies, we can investigate scenarios in which agents must collaboratively, through complementary action, arrive at a goal location known to only one agent.

While it is typically intractable to find the optimal solution to many multi-agent coordination problems, these problems are particularly important because their inherent challenges highlight several important features of social interaction and group dynamics that need to be studied:

1. **Non-stationary World**: Agents are constantly adapting to the statistics of their environment, including other agents. Since other agents do not have a fixed method of interacting with the world a priori, the world is inherently non-stationary (Buşoniu, Babuška, & Schutter, 2008).

2. **Non-independent Sampling**: An agent's own actions affect its incoming sensory information and this in turn affects the regularities it can extract from the world (Lungarella & Sporns, 2005). Motor activity and sensory information obtained from the environment are interdependent; the way we move in the world shapes our understanding of it and these patterns of data have structure.

3. **Distribution of Knowledge**: Not all agents in the world have access to the same information or capabilities. The social realm is comprised of more than just a set of identical individual problem solvers (Hutchins, 1995).

Another prominent research direction in studying multi-agent systems is determining "(h)ow to develop... problem solving protocols (information flow) that enable agents to share results and knowledge in a timely, effective manner" (Sen, 1997). It is important to understand how a group of individual agents reorganizes in functional ways that alter the flow of information; we need to understand "what information goes where and in what form" (Hutchins, 1995) and how these pathways change. This situation is complicated by the

fact that researchers in Cognitive Science hold different assumptions about the internal organization and external behavior of agents, which specifies the model elements, and this constrains the possible ways to reconfigure information flow. This situation can be rectified, however, by utilizing a common formalism for comparing and contrasting the consequences of different sets of assumptions.

In this paper, we utilize a formal framework, Decentralized Partially Observable Markov Decision Processes (Dec-POMDPs), to place our problem of interest into a larger set of multi-agent coordination problems in order to investigate coordination problems when agents have access to different amounts of information (Karnowski, accepted). We then discuss how several assumptions about agent architecture map into specific changes in the problem structure, demonstrating how we can vary our hypotheses by altering the components of the Dec-POMDP. Through the use of multi-agent Q-learning, we can demonstrate the speed with which agents reorganize themselves into stable patterns of behavior that allow them to coordinate their actions and achieve a joint goal. This reorganization brings differences in performance, however, based on the assumptions made about agent capabilities. We utilize mutual information to measure the changes in statistical dependencies among streams of information and to show how agents' behaviors respond to environmental regularities. We conclude by discussing how one problem formulation may provide insights into the study of the evolution of communication and future directions in this area.

Methods

Decentralized Partially Observable Markov Decision Processes (Dec-POMDPs)

Dec-POMDPs (D. Bernstein, Zilberstein, & Immerman, 2000) are a way to formalize multi-agent coordination problems. They provide a common structure that aids in the discussion of related problems and the development of solution techniques. While there exist other frameworks that tackle problems of agent coordination and problem solving (Dec-POMDP-COM, MTDP, and COM-MTDP with perfect recall), many of them have been shown to be formally equivalent (Seuken & Zilberstein, 2008). The reason for the variety is that the frameworks emphasize different features. For instance, while Dec-POMDPs and Dec-POMDP-COMs (Dec-POMDPs with communication) (Goldman, Allen, & Zilberstein, 2007) are formally equivalent, the former tends to focus on bodily coordination in physical space and the latter with problems that also involve symbolic coordination. In addition to communication, frameworks often contain assumptions about the representational capacities of their agents, providing agents with, for example, the ability to model the goals or actions of other agents (Claus & Boutilier, 1998). Providing a language for researchers in Cognitive Science to systematize problems in cooperative multi-agent interactions and make explicit their assumptions about individual architecture will allow for a thorough comparison of current models and the exploration of regions between models with different assumptions.

Formally, a Dec-POMDP can be defined by a tuple $\langle \{Ag\}, S, \{A\}, P, \{\Omega\}, O, R \rangle$, where $\{Ag\} = \{1, 2, \ldots, n\}$ is the set of agents, S is the possible states of the world, $\{A\} = \{A_1\} \times \{A_2\} \times \ldots \times \{A_n\}$ is the set of joint actions (with $a = (a_1, a_2, \ldots, a_n)$ being a joint action and action a_i is the action of agent i), P is the transition function (with $P(s'|s,a)$ being the transition to state s' given current state s and joint action a), $\{\Omega\}$ is the set of possible observations, O is the matrix that defines the probability of seeing observation o given state s, and $R = R(s,a,s')$ is the reward for taking the joint action a in state s and transitioning to state s'. The goal of solving a Dec-POMDP is to find a joint policy $\pi = \{\pi_1, \pi_2, \ldots, \pi_n\}$ (where each π_i is a local policy of one agent that maps an observation of a state to an action, i.e. $\pi_i : S \rightarrow A_i$) such that the group minimizes some cost function over time (similarly, it can maximize a reward function).

Multi-agent Q-learning

Dec-POMDPs are a useful abstraction which allows for a common language when speaking about coordination problems. These problems, are typically difficult to solve (D. Bernstein et al., 2000), but solution algorithms are a current research trend (Spaan & Oliehoek, 2008). Another way to address these problems is to use on-line adaptive heuristic algorithms that provide good approximate solutions, such as Q-learning (CJC, 1989), as they stochastically approximate off-line learning of optimal policies. In this paper, we use the Q-learning algorithm in a multi-agent context (Buşoniu et al., 2008). Within each agent, state-action pairs are strengthened depending on the outcome of the chosen action. For instance, if an agent transitions to state s' after performing action a while in state s, an agent will receive a reinforcement R and update the value of that state-action pair (s,a):

$$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha(R + \gamma max_{a' \in A} Q(s',a')) \quad (1)$$

Other parameters relate to the learning algorithm itself. The learning rate, α, determines the degree to which the current state is updated given new experience, and the discount factor, γ, specifies how influential future states and actions are to the current state. In this experiment, actions were chosen in a greedy manner.

Behavioral Info-Dynamics

Consider an isolated animal collective X consisting of n freely moving animals. Temporal data is collected on each animal's behavior generating a unique time series. Given a collection of sensorimotor time series data from a set of animals, we can measure statistical dependencies during different behavioral patterns. Tononi, Sporns, and Edelman (1994) (and later Tononi, Edelman, and Sporns (1998)) introduced a set of appropriately defined information-theoretic measures to capture the statistical properties of a system with n components. While their methods were originally designed to study

neural systems, more recent work has adapted these measures to study sensorimotor coordination in embodied agents by collecting sensor and motor time series data (Lungarella, Pegors, Bulwinkle, & Sporns, 2005). We utilize a Python implementation of these measures (available at `https://github.com/OpenCV-at-DCog-HCI/BID`) to further extend these measures to study the behavior of a system of agents. In this paper, we focus only on the mutual information between pairs of time series. Depending on their interaction with the world, solitary agents and collections of agents exploit different statistical dependencies among streams of information. We can show these changes by measuring mutual information (Sporns, Karnowski, & Lungarella, 2006; Di Prodi, Porr, & Wörgötter, 2010).

Entropy defines the uncertainty inherent in a time series, or the average amount of information present. For instance, if knowing the state of the system at a given point in time will give you a lot of information about the time series as a whole, then this will contribute to a lower entropy. This could happen if that state is highly unlikely, and thus is more informative. If every state, however, is equally likely, then knowing the state at one point in time gives no information about the time series as a whole and entropy is maximal.

$$H(X) = -\sum_{j=1}^{n} p(x_j) log(p(x_j)) \quad (2)$$

Mutual information measures the dependence between two distributions (and in our case, time series). It is defined as the Kullback-Leibler distance (D_{KL}) between the joint distribution $p(X_1,X_2)$ and the independent distribution $p(X_1)p(X_2)$. Mutual information is also defined as the sum of the entropies of the individual parts with the joint entropy subtracted out.

$$MI(X_1,X_2) = D_{KL}[p(X_1,X_2)||p(X_1)p(X_2)] = H(X_1)+H(X_2)-H(X_1,X_2) \quad (3)$$

Any dependence between the two time series will increase the mutual information between them. For instance, if the state of one agent provides a lot of information about the state of another agent, this will result in higher mutual information. If the agents are completely independent, then this predictive power is lost, and mutual information will be zero.

Problem and Experimental setup

To explore how two agents could coordinate via complementary actions to arrive at a hidden goal, we created an extension of the 'block pushing problem' (Matarić, 1996; Sen et al., 1994) where two agents are tasked to move from a start location to the goal, which is one of two possible locations, and follow as closely as possible a path P between the two. At every timestep, Agent i uses a force $\vec{F}_i$, where $0 \leq |\vec{F}_i| \leq F_{max}$ on the block at an angle θ_i, where $0 < \theta_i < \pi$, which results in the block being offset by $|\vec{F}_i|\cos(\theta)$ in the x direction and $|\vec{F}_i|\sin(\theta)$ in the y direction. The new position of the block is calculated by vector addition of the displacement created by the two agents. The new coordinates are then assigned to the correct discrete bin. The location of the block is used as feedback for the agents, depending on which scenario is being considered.

In our problem, $\{Ag\}$ is a set of two agents, S is the x-coordinate in a 20x20 grid world, the actions are a vector-addition of individual agent actions that combine force and angle ($0.2 \leq |\vec{F}_i| \leq 2.0$) in 0.2 increments and $15 \leq \theta_i \leq 165$ in 15 degree increments), P is deterministic (the probabilities of moving to the next state given a joint action is 1 and the rest are zero), the set of observations is always the current x-coordinate in the grid world but more information is added depending on the scenario (for the agent with the goal, the current goal is also added to the observation), Ω is deterministic (the probabilities of an agent perceiving a particular observation given a state is 1 and the rest are zero), and the feedback depends on the scenario.

The first goal of our study was to establish a baseline. We implemented the scenario as found in (Sen et al., 1994):

0. **Agent 2 also knows goal (Full Information):** Both agents receive an observation of their x-coordinate and the goal. Their feedback is a function of their distance from the goal path P.

Even though there are two possible paths, there is only one goal for each trial, and therefore our agents acted in similar manner and replicated the results obtained by Sen et al. (1994). We then set out to construct a situation where there is a disparity in the amount of information accessible to each agent. In our 'base case', we consider the impact of removing information about the goal from Agent 2 and only allowing Agent 1 to have this knowledge. From here, our models were motivated by research agendas within Cognitive Science. Given different assumptions of agent architecture, we alter the Dec-POMDP in specific ways:

1. **Agent 1 knows the goal but Agent 2 does not ('Base Case'):** Agent 1 remains identical to previous results, but the observation Agent 2 receives does not contain information about the goal. The feedback for Agent 2 is a function of the distance from the closest path (i.e. when there is no information about the goal, the closest path is the best)

2. **Agent 2 tracks probability of goal ('Theory of Mind'):** Giving an agent the ability to represent the goal of another agent and make inferences about that goal given data is one way to conceptualize Theory of Mind. In this situation, Agent 2 begins a trial with the prior belief that either goal is the possible target. At each time step, the state of the world is a sample with which Agent 2 updates its belief of the current goal via Bayes rule. The probability of this sample is the probability that the x-coordinate is sampled from a Gaussian distribution with the x-coordinate of the goal being the mean and a standard deviation of 2.5 (Altering this distribution is future work). The probability space was discretized into 10 bins. The feedback for Agent 2 is

an weighted average (given current belief) of the feedback for both paths.

3. **Agent 1 can make sounds ('Communication'):** Agent 1 produces either a 0 or 1 which becomes part of the state which Agent 2 will experience on the next time step. The feedback for Agent 2 is a function of the closest path.

4. **Agent 1 can make sounds and Agent 2 tracks probability of goal ('Theory of Mind' and 'Communication'):** This is a combination of the previous two alterations. The feedback for Agent 2 is the weighted average of the feedback for both paths.

The feedback in each of these cases is determined by a function of the distance from the desired path, $f(\delta x) = K * a^{-\delta x}$, similar to the original setup in Sen et al. (1994). This provides a high value for being on the path and an exponentially decreasing value further away from the desired path. Starting out the learning process with high values for state-action pairs and providing feedback after every trial was another feature in Sen et al. (1994) that allowed the agents to explore the available actions (alternatively, one could set the values in the beginning to be zero, but receiving feedback after just one trial would bias the agent to take the same path every trial). Also, any updates to state-action pairs could not be larger than the original high value (in our case, this was set to 100).

At the beginning of every trial, the two agents start at $(x,y) = (10,0)$ and the goal is randomly chosen from two options: $(3,20)$ or $(17,20)$. They make individual actions which combine into a joint action as outlined above. If the agents move the object outside of the 20x20 grid world, then the trial ends. Similarly, if the agents arrive at the goal state, the trials ceases. In the rare chance that agents would take more than 100 timesteps, the trial would also stop (forcing the angles to not allow agents to travel parallel to the x-axis helps alleviate this problem). An additional feature incorporated into the world dynamics was an automatic movement forward if the agents did not move forward enough on a trial. This was added to ensure agents did not remain still and allowed for better convergence.

Results

In our experiments, agents always began with equally valuable state-action pairs and this caused their actions to be selected randomly. Over many trials, as agents adjust the values of different actions within each state, their behaviors begin to become patterned. Practices reduce the entropy of the shared environment, which leads to better policies and to a decrease in the average distance from the goal path. One would suspect, however, that performance would be best when there is complete information for both agents and that scenarios in which one agent has partial and incomplete information, the resulting joint actions would lead to poorer performance. This is not what we find, as shown in Figure 1. Having the ability to produce and utilize sounds allows agents, over time, to perform better than those with complete information. Having the ability to represent and make inferences about the goals of another agent provides even more improvement in joint coordination.

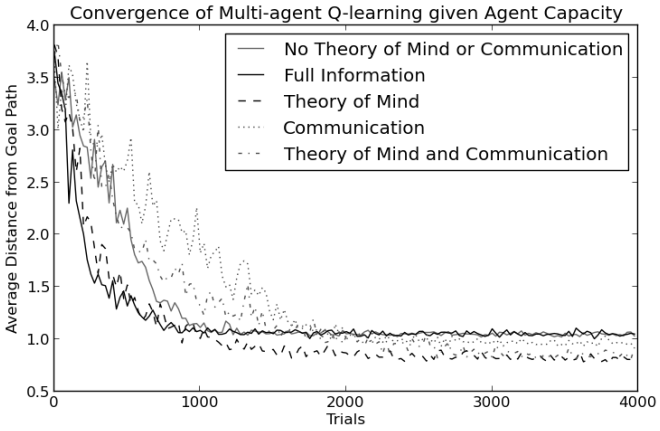

Figure 1: The average distance of the actual path from the goal path given different agent assumptions ($\alpha = 0.01$,$\gamma = 0.9$). Each experiment had 5000 trials and the data has been averaged over 100 experiments. Other learning rates ($\alpha \in \{0.1, 0.2, 0.3\}$) resulted in the similar patterns of performance with different rates of convergence.

We can determine how the two agents functionally reorganized themselves based on the levels of statistical dependence between different data streams. Mutual information provides a way to measure how predictable one data stream is from another. As we can see in Figure 2, both the scenario in which Agent 1 and Agent 2 have full knowledge of the goal and the 'base case', where Agent 2 does not know the goal, there is an increase in the mutual information between the x-coordinate and the angle of Agent 2 but this mutual informativeness plateaus. In the scenarios where there is Theory of Mind, Agent 2 is receiving a wealth of information about the goal through its current location but not necessarily needing to rely on any connection between its angle action choices and its location, which would have forced it to be more precise in its actions. In the scenarios with sound, there is a lot of extra structure in the shared environment that becomes highly predictive of the x-coordinate and therefore in the actions of Agent 2, including the angle. Another situation was created in which Agent 1 produced a sound but the state also included another random noise (to take away the special nature of the sound but not its ability to be manipulated). While the graph does not show the full increase of MI, other simulations showed this had the same trend as the case with communication, just over a longer period of time. This makes sense if agents were learning to utilize structure, but randomness was slowing this process down.

We did not find that the forces with which agents pushed the box had any predictive power for other data streams. When there was an increase in mutual information, it ap-

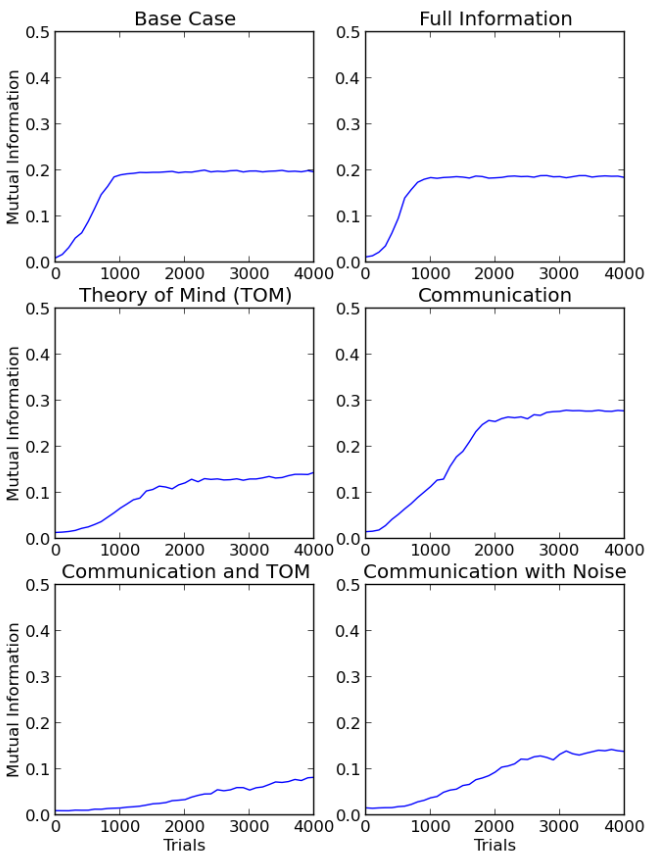

Figure 2: The mutual information between the x-coordinate and the angle of Agent 2.

peared to be due to the high predictability of angle and x-coordinate. As the world dynamics forced agents ahead one step if they did not apply enough force, it may have been the case that this affected the importance of force as a predictive element. This is probably not the case, however, as the agents in our model (and those in Sen et al. (1994)) only observe the x-coordinates, which would in turn dampen some of the informativeness of force in agent action choices.

Discussion

In this paper, we have discussed the benefits of utilizing a common theoretical framework for addressing cooperative multi-agent problems in Cognitive Science and demonstrated how changes to framework elements can encapsulate various hypotheses about agent actions and internal representational capacities. We have designed a new multi-agent problem, focusing on understanding the acquisition of complementary actions in a goal-directed task where there is an information disparity. We used Q-learning, an algorithm commonly used in modeling single agent decision making, in a multi-agent setting to investigate how agent hypotheses affect the convergence of the learning process. And finally, we used mutual information to quantify how informative one data stream, the x-coordinate, is about another data stream, the angle chosen by Agent 2 and charted the changes in this informativeness over time.

The results for this particular problem formulation provide a partial ranking of models based on performance. There are, however, a couple of caveats. First, while our simulated agents chose their actions in a greedy manner, different results might be obtained through other action selection methods, such as using a Boltzmann action selection mechanism. Second, Dec-POMDPs are typically used when there is some uncertainty in state transitions (due to modeling motor noise) or observations (due to sensory noise or partial view of the world). While this problem does not utilize this feature, future work manipulating these parameters may change the success of models with different assumptions about agent architecture.

This work highlights several of the open problems in the study of the emergence of communication, as it simultaneously investigates the origin of signaling channels, the sources of representation in signals, and the roles of social interaction in learned communication systems (Lyon, Nehaniv, & Cangelosi, 2006).

Future work related to this particular example will strive to explore how agents could learn to discover that one information stream is informative about another, a hallmark of communication. As a starting point, for instance, we are particularly interested in the case where the agents have an ability to put structure into the shared environment through sounds. In this case, it could be that the agent with the goal is able to create noises, which allows the second agent to adjust its policy given this external structure. This in turn forces more regular behavior to which the speaking agent can then adjust. Originally, the noise was not functionally related to the current state; in the beginning, sounds just happened. As engagement proceeds, that noise ends up carrying information, and at that moment, the sounds would become a signaling channel.

This process, however, hasn't held any commitments to the content of that signaling channel. It may turn out that the speaking agent, through features of the algorithm, converges on highly rewarding action-sound pairings and the second agent only need adjust its behavior accordingly. In either case, we suspect that putting structure out into the world may create stable regularities with which agents could take advantage and eventually internalize (Vygotsky, 1978). Agent interactions themselves would be the determining factor behind the sources of representations in the signals they employ. In problems similar to ours, it is often the case that multi-agent Q-learning fails, precisely because neither agent experiences a stationary environment (Claus & Boutilier, 1998). Placing stationary-creating behavior at the center of new algorithms is also possible future work.

Here we have shown that we can operationalize several assumptions in Cognitive Science and discover what structure and organization emerge from these hypotheses. In the present examples, however, agents are endowed with cer-

tain abilities a priori. We would really like to explore the conditions under which language-like abilities and Theory of Mind-like processes could emerge from ongoing interactions between autonomous agents. Additional future work will look at the space between these hypotheses and how various learning algorithms could take agents from a lack of abilities to a state where additional mental abilities have emerged through agent interactions.

Acknowledgments

The authors would like to thank Chris Johnson, Ben Bergen, and Ben Cipollini for helpful discussions. Jeremy Karnowski is a Jacobs Fellow and is the recipient of a CARTA Graduate Fellowship in Anthropogeny.

References

Bernstein, D., Zilberstein, S., & Immerman, N. (2000). The complexity of decentralized control of markov decision processes. In *Proceedings of the sixteenth conference on uncertainty in artificial intelligence* (pp. 32–37).

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

Buşoniu, L., Babuška, R., & Schutter, B. De. (2008, March). A comprehensive survey of multi-agent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *38*(2), 156–172.

CJC, H. (1989). Learning from delayed rewards. *Cambridge University, Cambridge, England, Doctoral thesis*.

Claus, C., & Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. In *In proceedings of national conference on artificial intelligence (aaai-98* (pp. 746–752).

Di Prodi, P., Porr, B., & Wörgötter, F. (2010). A novel information measure for predictive learning in a social system setting. *From Animals to Animats 11*, 511–522.

Goldman, C., Allen, M., & Zilberstein, S. (2007). Learning to communicate in a decentralized environment. *Autonomous Agents and Multi-Agent Systems*, *15*(1), 47–90.

Hazlehurst, B., & Hutchins, E. (1998). The emergence of propositions from the co-ordination of talk and action in a shared world. *Language and Cognitive Processes*, *13*(2-3), 373–424.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Hutchins, E., & Johnson, C. (2009). Modeling the emergence of language as an embodied collective cognitive activity. *Topics in Cognitive Science*, *1*(3), 523–546.

Karnowski, J. (accepted). Modeling collaborative coordination requires anthropological insights. *Topics in Cognitive Science*.

Lungarella, M., Pegors, T., Bulwinkle, D., & Sporns, O. (2005). Methods for quantifying the informational structure of sensory and motor data. *Neuroinformatics*, *3*(3), 243–262.

Lungarella, M., & Sporns, O. (2005). Information self-structuring: Key principle for learning and development. *Proceedings The 4nd International Conference on Development and Learning 2005*, 25–30.

Lyon, C., Nehaniv, C., & Cangelosi, A. (2006). *Emergence of communication and language*. Springer.

Matarić, M. (1996). Learning in multi-robot systems. *Adaption and Learning in Multi-Agent Systems*, 152–163.

Sen, S. (1997). Multiagent systems: Milestones and new horizons. *Trends in Cognitive Sciences*, *1*(9), 334–340.

Sen, S., Sekaran, M., Hale, J., et al. (1994). Learning to coordinate without sharing information. In *Proceedings of the national conference on artificial intelligence* (pp. 426–426).

Seuken, S., & Zilberstein, S. (2008). Formal models and algorithms for decentralized decision making under uncertainty. *Autonomous Agents and Multi-Agent Systems*, *17*(2), 190–250.

Spaan, M., & Oliehoek, F. (2008). The multiagent decision process toolbox: Software for decision-theoretic planning in multiagent-systems.

Sporns, O., Karnowski, J., & Lungarella, M. (2006). Mapping causal relations in sensorimotor networks. In *Proc. of the 5th int. conf. on development and learning.*

Tononi, G., Edelman, G., & Sporns, O. (1998). Complexity and coherency: integrating information in the brain. *Trends in cognitive sciences*, *2*(12), 474–484.

Tononi, G., Sporns, O., & Edelman, G. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(11), 5033–5037.

Vygotsky, L. (1978). *Mind in society*. Harvard University Press.