Challenges for Multi-Agent Coordination Theory Based on Empirical Observations

Victor Lesser and Daniel Corkill
School of Computer Science
University of Massachusetts Amherst
Amherst, Massachusetts 01003
{lesser,corkill}@cs.umass.edu

ABSTRACT

Significant research progress and understanding about the nature of coordination has been made over the years. Development of the DCOP and DEC-MDP frameworks in the past decade has been especially important. Although these advances are very important for multi-agent coordination theory, they overlook a set of coordination behaviors and phenomena that have been observed empirically by many researchers since the early years of the field. The goal of this paper is to challenge researchers in multi-agent coordination to develop a comprehensive formal framework that explains these empirical observations.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*

Keywords

Coordination, Agent Cooperation, Distributed Problem Solving

1. INTRODUCTION

The study of coordination and cooperation among agents has been at the heart of the multi-agent field since its inception [11, 5]. Since this early work, significant research progress and understanding about the nature of coordination has been made [8, 9, 19, 13, 10, 20]. Especially important has been the development of distributed constraint optimization (DCOP) [24] and decentralized Markov decision processes (DEC-MDPs) [1] frameworks over the last decade. These formal frameworks allow researchers to understand not only the inherent computational complexity of coordination problems, but also how to build optimal or near-optimal coordination strategies for a wide variety of multi-agent applications.

These directions are very important for multi-agent coordination theory but overlook a set of coordination behaviors and phenomena that have been observed empirically by researchers since the early years of the field. These behaviors have often been exploited by researchers building efficient heuristic coordination mechanisms, but rarely are they understood deeply or explained formally.[1] Exploiting these phenomena usually requires taking a more statistical view of coordination behavior and taking into consideration the underlying distributed search process being coordinated. This is in contrast with current formal approaches that look for some explicit structural interaction pattern associated with a problem description that reduces computational complexity.

The goal of this paper is to challenge researchers in multi-agent coordination to develop a comprehensive formal framework that explains these empirical observations. A deeper, formal understanding of these phenomena could help researchers develop new and more efficient coordination strategies—possibly similar to how the study of phase transitions in NP-hard problems [16] opened up new perspectives to researchers studying computational complexity and search mechanisms.

In the remainder of this paper, we will discuss the following coordination phenomena that need to be better understood and, ideally, explained more formally: 1) that structural interrelationships among agent activities inherent in the problem description are not necessarily indicative of the communication complexity necessary for effectively coordinating agents, 2) that modifying local problem solving to make it more predictable or less responsive or decreasing the frequency of coordination sometimes improves agent coordination, 3) that a greedy and incremental approach to coordination can be an appropriate heuristic, and 4) that dynamic adaptation of a coordination strategy to the current state of the network problem solving can be effective.

2. COORDINATION AND INTERACTION

Most of the formal research in coordination theory uses the explicit structural patterns of interaction among agents in reducing the computational effort required to find an optimal strategy. These structural relationships are typically obtained from characteristics of the problem description. An early example of this approach was work on transition independent DEC-MDPs [15], where actions taken in one agent do not affect the outcome of actions taken by another agent. However, we hypothesize that the existence of these structural relationships does not always indicate the communication requirements necessary for implementing an effective coordination strategy. An example of this was the observation by Mostafal that, for at least one class

[1] There are some exceptions [6, 17], however those are limited to specific and narrowly defined coordination problems.

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

of problems with structural interaction patterns that (on the surface) indicated that explicit communication of agent states would be advantageous, it was very hard to find specific problem instances of this class where the optimal coordination strategy actually required communication [14]. A slightly different but related observation was made in the early work on solving DEC-MDPs [22]. The approach they used first solved a centralized version of the coordination problem framed as a Multi-agent Markov Decision Process (MMDP). In this MMDP solution, agents were aware of the state of other agents, which implied that the distributed implementation of each agent's policy required the agent to communicate its current state to other agents at each time step. Through analysis of this optimal centralized policy, it was shown that many of these communication actions were unnecessary, and that an optimal coordination policy could still be maintained in at least two-agent examples. From our perspective, what was even more interesting occurred when approximations were introduced to this optimal policy derived from the MMDP solution by assuming, from a statistical perspective, what were the likely problem-solving states of the other agent given their local control policies. With these heuristics, they demonstrated that they could eliminate large amounts of communication with only a slight reduction in optimality.

Zhang and Lesser achieved similar results in their recent work on multi-agent reinforcement learning [27] where they used a DCOP algorithm to coordinate agent learning to approximate a centralized learning algorithm. In this case, they realized that instead of having one massive DCOP that spanned all agents, they could break the DCOP into a set of much smaller independent DCOPs, which significantly reduced the amount of communication required to implement the coordinated learning, with only a slight reduction in the utility of the learned policies of the agents. In developing this dynamic decomposition of the DCOP, they used a statistical view of agents' states based on their current policy to find situations in which not knowing the current states of specific agents would not significantly decrease other agents' utility. Again, the need for communication among agents did not always relate directly to structural interaction in the problem description, especially when a slight decrease in overall utility was acceptable. We hypothesize that there is something going on that has not been modeled by the explicit structural relationships on agent activities as defined by the problem description. It is not the existence of all interaction relationships that needs to be modeled, but something more nuanced where a trade-off between optimality and communication can be expressed. A theory is needed that connects the characteristics of the problem description and the character of optimal or near-optimal coordination strategies. When only the agents' key interactions are considered, the agents often appear significantly more loosely connected (more nearly decomposable [18]) than would be expected by the existence of all structural interactions among agents.

There is also another intuition present in these examples. The constraints coming from the local ordering of agent activities are exploited implicitly by the coordination strategy to reduce the need for explicit coordination among agents. To the degree that there is flexibility in how to organize local problem solving, it becomes more likely that the coordination strategy can find a combined ordering of local agent activities that reduces the need for explicit coordination among agents. From this perspective, the introduction of non-optimal local behavior, if done astutely, can present new options for finding combined agent activity orderings; thus, potentially reducing coordination overhead.

3. COORDINATION AND COMPUTATION

Another way of thinking about the observations in Section 2 is in terms of what assumptions one agent can make about the state of other agents with whom they potentially interact. In learning theory, this idea is discussed in terms of the concept of a non-stationary environment: the more non-stationary the environment is, the harder the learning. Thus, coordination techniques that can decrease or change the nature of the non-stationary environment in multi-agent learning caused by concurrent learning in neighboring agents can improve learning performance significantly in terms of both the speed of convergence and the likelihood that convergence will actually occur. The basic approach developed by the multi-agent reinforcement community is to make the local agent learning algorithm change in slower and/or more predictable ways [2, 26, 28]. In this way, even though individual agent learning may not be as efficient from a local perspective, learning from a system-wide perspective can converge more quickly and to better solutions.

This issue of a non-stationary environment also occurred in different guises in earlier work on developing both heuristic and formal coordination strategies. These examples have an interesting connection with the multi-agent reinforcement learning example discussed above: they all involve the use of iterative algorithms, where the same basic process is repeated on each cycle as new information is received. Brooks and Lesser coined the term "simultaneous-update uncertainty" to describe this non-stationary environment characteristic [3]. They worked on the problem of distributed traffic light control using an iterative algorithm and developed such techniques as modulating the magnitude of changes on any cycle, giving priority to certain neighboring traffic light agents' information changes over other agents' information, and modulating the frequency of updates based on the state of the agents' current traffic control pattern. All of these strategies decreased simultaneous-update uncertainty and improved performance. Similarly, Fernandez, et al., found that the "active introduction of message delays by agents can improve performance and robustness while reducing the overall network load" for distributed constraint satisfaction algorithms (DCSP) [4].

Our intuition for explaining this behavior relates to how an iterative improvement search process works. If the search is started with a tentative solution that is partially correct, performance improves significantly. However, even without a good starting point, this type of search can still be effective because it can often find tentative solutions quickly that contain fragments/partial-solutions that correspond to fragments of the correct solution. These fragments direct the search process to find the correct solution. For example, consider a distributed search such as asynchronous weak-commitment search (AWC) [23] used by Fernandez, et al. [4], where each agent is solving a component of the overall problem. If the coordination does not allow clusters of agents to form consistent solutions with sufficient frequency due to stationarity issues, then the distributed search will take much longer in the case of complete algorithms such as AWC and, in the case of incomplete algorithms, lead to oscillation or convergence to suboptimal solutions. Thus, by slowing down the frequency of updates, it is more likely that groups of agents will construct consistent fragments of the overall solution; this will in turn speed up the overall search process. Another way of framing this is that all these "heuristic approaches are intended to reduce the oscillation during search caused by concurrent learning or local partial solution update. In some sense, they serialize or coordinate agents' local search activities to improve the performance" [25].

Unfortunately, to our knowledge a formal theory of distributed search that explains in detail why the above heuristic approaches to improving agent coordination work has not been developed. What is missing is a theory that explains how both the character and frequency of incorrect or out-of-date information affects the performance of a coordination strategy and, ultimately, overall network problem solving. The difficulty in developing such a formal theory is that the consequences of this inaccurate information and associated problem solving are not confined to individual agents but can propagate throughout the agent network. Thus, there is a need to incorporate some type of statistical model of the distributed search process being coordinated and its associated intermediate states into a formal analysis framework for explaining these phenomena.

4. COORDINATION AND ENVIRONMENT

The issue of moderating the frequency of coordination has also come in another guise. Durfee and Lesser introduced the trade-off between predictability and responsiveness, where communication and computation costs associated with coordination are modulated by varying the conditions when an agent indicates that its current state does not match the expectations used in the current coordination strategy [7]. In this case, a wider tolerance for variance from the expected agent behavior leads to more predictability in a coordination strategy (since it is less likely to be revised) with the consequence that the strategy was not as responsive to the details of the current agents' states and thus the coordination was not as precise. However, they observed that, given the additional costs and delays of being highly responsive, it may be better to use a less responsive coordination strategy. Further, they found in one case that even if these additional costs were discounted, it remained better to be more predictable because the coordination strategy was constantly changing on each cycle, causing unnecessary backtracking of agent problem solving in a way similar to what we discussed previously in Section 3.

More generally, depending on the statistics of the environmental conditions in terms of resource availability, task loading, and predictability of task behavioral characteristics, very different coordination strategies are appropriate. Without going into detail, here are our summaries of some of the observations. The first observation is that in environments where there is very high or very low task loading or high variance in agent behavior, simple coordination strategies work quite well.[2] It is only in situations where there is a "sufficient" level of predictability about agent behavior or intermediate levels of task loading that complex coordination strategies are advantageous. This last point relates to the nature of phase transition, where the difficulty of solving problems increases significantly around the phase transition. We suggest that there are similar phase transitions going on in agent coordination and that it is in those transition regions where more complex control is advantageous. The second observation is that even though DCOPs can be used to build an optimal coordination strategy, simpler non-optimal heuristic coordination strategies that only consider the major interactions among agents and do not deal with contingencies directly (or deal with them in only limited ways) do quite well in most coordinating situations (see [20, 21]). These approaches re-coordinate when necessary based on the actual contingent event rather than attempt to prevent such situations from occuring. The intuition behind this observation is again that agent interactions in most situations are more loosely connected than would be expected, and most incorrect coordination decisions can be tolerated without severe harm to overall agent performance. The question for us is whether there is a formal way from looking at a problem and its environmental description to understand what contingencies, agent activity horizons, and problem-solving states of other agents need to be considered in order for coordination to work effectively.

5. COORDINATION DYNAMICS

This final section deals with the dynamics of coordination. We have discussed above that, depending on the environmental characteristics, very different approaches to coordination are appropriate. The same holds for the underlying distributed search or learning process being coordinated. In this case, the character of the distributed search process for a single problem may vary significantly over its lifetime. Mailler and Lesser recognized this in developing a very effective approach to distributed constraint satisfaction where the scope of control (partial centralization of control) varied on each cycle based on the current constraint interactions among the partial solutions constructed at different agents [12]. Similarly, Zhang and Lesser used a strategy for coordinating multi-agent learners that dynamically changed the scope of control based on the strength of interaction among current learned policies of agents [27, 28]. More generally, a coordination strategy that can adapt to the current situation seems crucial where the environment or network problem solving is evolving dynamically and rapidly, and different situations require different approaches to coordination.

6. CONCLUSIONS

Our intuition is that all these experimental behaviors and phenomena are interrelated and that an integrated and formal treatment of them by future generations of researchers will lead to a much deeper understanding of the nature of coordination and cooperation and, more generally, decentralized control. Lacking from current formal frameworks are: 1) a statistical model of the underlying distributed search

[2] Decker showed formally, for a specific task allocation problem, that if there was high variance in the number of tasks associated with different agents more sophisticated coordination strategies that exploited meta-level control information did better [6]. However, this does not contradict our basic point because the specific instances of high variance in this case could be ascertained before the coordination strategy was constructed (based on meta-level information) rather than needing to be recognized during the execution of the coordination strategy.

process (network problem solving) that is being coordinated and its associated intermediate states and 2) formal treatment of concepts such as "near decomposability" and "satisfiability" developed by Simon for understanding effective coordination [18]. That is our challenge to the multi-agent field.

7. ACKNOWLEDGMENTS

This material is based in part upon work supported by the National Science Foundation under Award Numbers IIS-0964590 and IIS-1116078. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

8. REFERENCES

[1] D. S. Bernstein, S. Zilberstein, and N. Immerman. The complexity of decentralized control of Markov decision processes. In *Proceedings of the 16th International Conference on Uncertainty in Artificial Intelligence*, pages 32–37, Stanford, California, July 2000.

[2] M. Bowling and M. Veloso. Multiagent learning using a variable learning rate. *Artificial Intelligence*, 136:215–250, Apr. 2002.

[3] R. S. Brooks and V. R. Lesser. Distributed problem solving using iterative refinement. Technical Report 79-14, Department of Computer and Information Science, University of Massachusetts Amherst, Amherst, Massachusetts 01003, May 1979.

[4] Cesar Fernandez, et al. Communication and computation in distributed csp algorithms. In V. Lesser, C. L. Ortiz, Jr., and M. Tambe, editors, *Distributed Sensor Networks: A Multiagent Perspective*, chapter 12, pages 299–318. Kluwer Academic Publishers, 2003.

[5] R. Davis and R. G. Smith. Negotiation as a metaphor for distributed problem solving. *Artificial Intelligence*, pages 63–109, 1983.

[6] K. Decker and V. Lesser. An approach to analyzing the need for meta-level communication. In *IJCAI 1993*, pages 360–366, Chambery, France, Aug. 1993.

[7] E. Durfee and V. R. Lesser. Predictability versus responsiveness: Coordinating problem solvers in dynamic domains. In *AAAI 1988*, pages 66–71, St. Paul, Minnesota, Aug. 1988.

[8] E. H. Durfee and V. R. Lesser. Partial global planning: A coordination framework for distributed hypothesis formation. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-21(5):1167–1183, May 1991.

[9] N. R. Jennings. Commitments and conventions: The foundation of coordination in multi-agent systems. *The Knowledge Engineering Review*, 8(3):223–250, 1993.

[10] V. R. Lesser. Reflections on the nature of multi-agent coordination and its implications for an agent architecture. *Autonomous Agents and Multi-Agent Systems*, 1(1):89–111, Mar. 1998.

[11] V. R. Lesser and D. D. Corkill. Functionally accurate, cooperative distributed systems. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-11(1):81–96, Jan. 1981.

[12] R. Mailler and V. R. Lesser. Asynchronous partial overlay: A new algorithm for solving distributed constraint satisfaction problems. *Journal of Artificial Intelligence Research*, 25:529–576, Jan–Apr 2006.

[13] Makoto Yokoo, et al. The distributed constraint satisfaction problem: Formalization and algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 10(5):673–685, Sep/Oct 1998.

[14] H. Mostafa. Private communications on the difficulty of finding optimal policies that benefit from explicit communication for the Mars Rover scenarios described in her Ph.D. dissertation, 2011.

[15] R. Becker, et al. Solving transition independent decentralized Markov decision processes. *Journal of Artificial Intelligence Research*, 22:423–455, Jul–Dec 2004.

[16] Rémi Monasson, et al. Determining computational complexity from characteristic 'phase transitions'. *Nature*, 400:133–137, July 1999.

[17] S. Sen and E. H. Durfee. A formal study of distributed meeting scheduling. *Group Decision and Negotiation*, 7(3):265–289, May 1998.

[18] H. A. Simon. *The Sciences of the Artificial*. MIT Press, 1969.

[19] M. Tambe. Towards flexible teamwork. *Journal of Artificial Intelligence Research*, 7:83–124, Jul–Dec 1997.

[20] V. R. Lesser, et al. Evolution of the GPGP/TÆMS domain-independent coordination framework. *Autonomous Agents and Multi-Agent Systems*, 9(1):87–143, July 2004.

[21] T. Wagner, V. Guralnik, and J. Phelps. A key-based coordination algorithm for dynamic readiness and repair service coordination. In *AAMAS 2003*, pages 757–764, Melbourne, Australia, July 2003.

[22] P. Xuan and V. Lesser. Multi-agent policies: From centralized ones to decentralized ones. In *AAMAS 2002*, Bologna, Italy, July 2002.

[23] M. Yokoo. Asynchronous weak-commitment search for solving distributed constraint satisfaction problems. In *Proceedings of the First International Conference on Principles and Practice of Constraint Programming (CP '95)*, pages 88–102, Cassis, France, Sept. 1995.

[24] M. Yokoo and E. H. Durfee. Distributed constraint optimization as a formal model of partially adversarial cooperation. Technical Report CSE-TR-101-91, University of Michigan, 1991.

[25] C. Zhang. Private communication, 2013.

[26] C. Zhang and V. Lesser. Multi-agent learning with policy prediction. In *AAAI 2010*, pages 927–934, Atlanta, Georgia, July 2010.

[27] C. Zhang and V. Lesser. Coordinating multi-agent reinforcement learning with limited communication. In *AAMAS 2013*, pages 1101–1108, St. Paul, Minnesota, May 2013.

[28] C. Zhang, V. Lesser, and S. Abdallah. Self-organization for coordinating decentralized reinforcement learning. In *AAMAS 2010*, pages 739–746, Toronto, Canada, May 2010.