

# Run-Time Selection of Coordination Mechanisms in Multi-Agent Systems

Rachel A. Bourne<sup>1</sup> Cora B. Excelente-Toledo<sup>2</sup> and Nicholas R. Jennings<sup>3</sup>

**Abstract.** This paper presents a framework that enables autonomous agents to dynamically select the mechanism they employ in order to coordinate their inter-related activities. Adopting this framework means coordination mechanisms move from the realm of being imposed upon the system at design time, to something that the agents select at run-time in order to fit their prevailing circumstances and their current coordination needs. Empirical analysis is used to evaluate the effect of various design alternatives for the agent's decision making mechanisms and for the coordination mechanisms themselves.

## 1 INTRODUCTION

Effective coordination is essential if autonomous agents are to achieve their goals in multiple agent systems. Such coordination is required to manage the various forms of dependency that naturally occur when the agents have inter-linked objectives, when they share a common environment, or when there are shared resources. To this end, a variety of mechanisms have been developed to address the coordination problem. At one end of the spectrum, there are social laws [8] which are laid down, long-term rules that prescribe how to behave in a given society (e.g., drive on the right-hand side of the road) such that if all the agents adhere to the rules then coordination should ensue. At the other extreme, there are a number of one-shot protocols (e.g., the Contract Net [9]) that coordinate the short-term activities of agents in order to accomplish a specific task. In-between are mechanisms, such as Partial Global Planning [3], that enable the agents to exchange high-level objectives in order to try and ensure coordination over a range of activities on a medium-term time horizon.

All of these *coordination mechanisms* have different properties and characteristics and are suited to different types of tasks and environments. They vary in the degree to which coordination is prescribed at design time, the amount of time and effort they require to set up a given coordination episode at run-time, and the degree to which they are likely to be successful and produce coordinated behaviour in a given situation. In the majority of cases, these dimensions act as forces in opposing directions. As observed by contingency theory [5], coordination mechanisms that are guaranteed to succeed typically have high set up and maintenance costs, whereas mechanisms that have lower set up costs are also more likely to fail. In short, there is no universally best coordination mechanism.

<sup>1</sup> Dept. Electronic Engineering, Queen Mary and Westfield College, University of London, London, E1 4NS, UK, email: r.a.bourne@elec.qmw.ac.uk

<sup>2</sup> Dept. of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK, email: cbet99r@ecs.soton.ac.uk

<sup>3</sup> Dept. of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK, email: nrj@ecs.soton.ac.uk

In all of these cases, however, the choice of coordination mechanism is something that the designer imposes upon the system at design time. While this may be sufficient for predictable and stable environments, it is inappropriate in dynamic and open contexts because there is no scope for changing or modifying the mechanism to ensure there is a good fit with the prevailing circumstances. In such environments, it is important that the agents have a variety of coordination mechanisms, with varying properties, at their disposal and that they can then select the mechanism which is most appropriate for the task at hand. Thus for particularly important tasks, the agents may choose to adopt a coordination mechanism that is highly likely to succeed, but which will invariably have a correspondingly large set up cost. Whereas for less important tasks, a mechanism that is less likely to succeed, but which has lower set up costs, may be more appropriate.

Our long-term aim in this research is to develop agents that can reason about the process of coordination and then select mechanisms that are appropriate to their current situation. That is, the choice of coordination mechanism is selected at run-time by the agents that need to coordinate. This work advances the state of the art in three main ways. Firstly, the very idea of letting the agents dynamically select the coordination mechanism is an issue that has not explicitly been addressed within the field of multi-agent systems to date. Secondly, as well as identifying this new area, we present a formal framework for capturing the reasoning processes the agents require in order to perform in such a manner. Finally, we provide an empirical evaluation of the effect of such reasoning on both the individual agents concerned and on the overall system.

The remainder of this paper is structured in the following manner. Section 2 describes the scenario we use to explore issues related to the dynamic selection of coordination mechanisms. Section 3 details the agents' decision making processes and section 4 evaluates them empirically. Section 5 relates our work to that of others and section 6 summarises our initial findings and highlights the avenues of further research.

## 2 THE COORDINATION SCENARIO

Our exemplar domain takes the form of a grid-world in which some number of autonomous agents ( $A_i$ ) perform tasks for which they receive units of *reward* ( $R_i$ ). Each agent has a *specific task* ( $ST_i$ ) which only it can perform; there are other tasks which require several agents to perform them, called *cooperative tasks* ( $CTs$ ). Each task has a reward associated with it. Generally, the rewards for the  $CTs$  are higher than those for  $STs$  since they must be divided between the coordinating agents. An example of a typical grid at one instant in time with two agents, two  $STs$  and one  $CT$  is given in figure 1.

The agents move about the grid one step at a time, up, down, left

		$x \rightarrow$				
		1	2	3	4	5
$y \downarrow$	1	$A_2$				
	2		$CT$			
3					$A_1$	
4		$ST_1$				
5				$ST_2$		

**Figure 1.** A typical coordination world grid.

or right, or stay still. At any one time, each agent has a single *goal*, either its ST or a CT over which it is coordinating. On arrival at a square containing its goal, the agent receives the associated reward. In the case of STs, a new one appears, randomly, somewhere in the grid, visible only to the appropriate agent. In the case of CTs, a new one appears, randomly, somewhere in the grid, but this is only visible to an agent who subsequently arrives at that square.

If an agent encounters a CT, en route to its current goal (i.e., its ST), it takes charge of the CT<sup>4</sup> and must decide on both whether to initiate coordination with other agents over this task, and if so which coordination mechanism (CM) it should use. In this context, each agent has a predefined range of CMs at its disposal. Each CM is parameterised by two key attributes: set up cost (in terms of time-steps) and its chances of success. For example, a CM may take  $t$  time-steps to set up (modelled by the agent waiting that number of time-steps before requesting bids from other agents) and have a probability,  $p$ , of success (thus when the other agent(s) arrive at the CT square, the reward will be allocated with probability  $p$ , with zero reward otherwise). An agent may well decide that attempting to coordinate is not a viable option, in which case it adopts the null CM (meaning the agent rejects adopting the CT as its goal).

The agent-in-charge (AiC) of the coordination selects a CM and, after waiting for the set up period, broadcasts a request for other agents to engage in coordination. The other agents respond with bids composed of the amount of reward they would require in order to participate in the CT and how many time-steps away from the CT square they are situated. Figure 2 gives the protocol the agents follow at each time-step; it highlights the specific decisions which have to be made (see section 3).

This initial presentation involves several simplifying assumptions; in particular common knowledge, a deterministic environment and straightforward coordination mechanisms. However, the framework is also intended to be flexible so that these and other assumptions will be relaxed in future work.

### 3 DECISION-MAKING PROCEDURES

This section formalises the decision procedures of the agents. To study the average impact of coordination mechanisms, an infinite horizon model of decision making [4] was adopted because we are interested in the long-term performance of agents; a finite horizon model may lead to erratic behaviour as the last time-step approaches. However, there are still two ways to model the agents' decisions: by using average reward per unit time or by discounting future rewards; the former was chosen, since it simplifies the decision analysis.

<sup>4</sup> If several agents arrive at a CT square at the same time, one of them is arbitrarily deemed to be in charge.

- [1] Each agent arrives at a square. If its goal is attained it receives reward and updates its goal.
- [2] If the agent finds a CT at the square it becomes AiC and decides to use CM =  $(t, p)$ , say. If  $t > 0$  it must wait  $t$  time-steps then broadcasts a request for coordination.
- [3] If a request for coordination is received, each agent decides on and submits its bid; if successful it adopts CT as new goal as does the AiC. If no suitable bids arrive, AiC returns to step [2].
- [4] Each agent decides on its next move according to its current goal and all agents make their move simultaneously.

**Figure 2.** Protocol for agents at each time-step.

All the agents in this scenario are assumed to be essentially cooperative though self-interested. However, to model the fact that this may not always be the case, we introduce the concept of an agent's *willingness to cooperate* (WtC); this factor,  $\omega$ , represents the weight an agent puts on opting to cooperate rather than collect its usual reward. When reward units, effectively the agent's utility, are of equal currency, a *neutral* agent only needs to receive the same reward from a CT as it would from its ST; this corresponds to a WtC factor of  $\omega = 1$ . The decision procedures described in this section will assume that agents are neutral but will include  $\omega$  to indicate where this factor comes into the calculations.

There are four types of decisions that agents are required to make: the direction to move in; which CM to adopt, if any; how much to bid when a request for coordination is received; and, how to determine which bid to accept, if any.

#### 3.1 Deciding on the direction of movement

An agent always has a target square in which its current goal is located. The agent decides to move towards its goal by selecting the direction, up, down, left, or right, probabilistically according to the ratio of up/down to left/right squares away from the goal it is.

For example, if an agent's ST is located 4 squares up and 3 squares to its right, then with probability  $\frac{4}{7}$  it will move up, and with probability  $\frac{3}{7}$  it will move right.

#### 3.2 Deciding which CM to use

An agent which, en route to its ST, encounters a CT, must decide whether to initiate coordination with another agent in order to perform it. To do this, the agent must determine whether there is any advantage in so doing. This depends not only on the reward that is being offered, but also on the CMs available, as well as various environmental factors which affect the expected demands of the potential coordinating agents.

To model the *expected demands* of the other agents, the AiC assumes that they are randomly distributed throughout the grid, and that their STs are similarly distributed. Thus some agents may be near the CT while others may be far away; likewise, for some agents there would be a significant deviation from their ST to reach to CT, while others may be able to coordinate over the CT en route to their

own STs. The AiC assesses the possible CMs on the basis of how long before the task can be performed (including both the set up time and the average distance away each agent is situated), and of how much reward it is likely to obtain after deducting the expected reward requirement of the other agent (based on the amount of time it must spend deviating from its path and the probability of success of the CM).

The AiC uses all these factors to assess each CM in terms of the amount of surplus reward it can expect, over and above what it expects to obtain during its normal course of operation, i.e., its own average reward per time-step,  $r$ . The AiC selects the CM which will maximise this surplus reward. Though this may not be a globally optimal criterion for deciding on which CM to use, it makes sense from a self-interested agent's point of view.

To formalise this decision procedure, consider an  $M \times N$  grid with reward size  $S$  for STs, and  $R$  for CTs, a coordination mechanism,  $CM = (t, p)$ , which costs  $t$  time-steps to set up and has a probability of success  $p$ . In this grid-world of known size, the agent can calculate the expected average distance (`ave_dist`) away of any randomly situated agent from the CT square, as well as the likely average deviation (`ave_dev`) such agents would have to make to get there. The AiC further assumes that all agents have similar average rewards and WtC factors as its own. Though in reality such common knowledge may not always be available, an agent may be able to build up a picture of its environment through past experience—clearly it needs some means of handling the uncertainty and this assumption is not an unreasonable preliminary approximation. Thus all agents use the same average reward per time step:  $r = \frac{S}{\text{ave\_dist}}$ .

Based on these figures, the agent can assess the average surplus reward from coordinating over the CT at  $[x, y]$  using  $CM_j = (t_j, p_j)$ . First, it must estimate its own cost in terms of how long the CM will take to set up and how long it expects to wait for another agent to arrive. Since the AiC would usually expect to receive  $r$  reward units per time-step, the cost of  $CM_j$  is given by:

$$\text{cost}_j(x, y) = r \times (t_j + \text{ave\_dist}(x, y))$$

Second, the AiC must estimate the average amount of reward the other agents will require:

$$\text{ave\_bid}_j(x, y) = \frac{r \times \omega \times (\text{ave\_dev}(x, y))}{p_j}$$

Using these estimates, the AiC can evaluate the expected surplus reward of  $CM_j$ :

$$\text{ave\_surplus}_j(x, y) = p_j \times R - (\text{cost}_j(x, y) + \text{ave\_bid}_j(x, y))$$

Note that the null CM is defined to have zero surplus.

When deciding which CM to adopt, the agent computes its expected surplus reward from each of them and selects that CM which maximises this value. If the surplus associated with all CMs is negative, the agent adopts the outside option of the null CM.

By means of illustration consider the scenario of figure 1. An agent occupies a  $5 \times 5$  grid and finds a CT with  $R = 6$  at square  $[3, 2]$ . Assume all agents have a WtC factor of  $\omega = 1$ . The average distance of other agents from  $[3, 2]$  is 2.6. Since the average distance between two random squares is 3.2, the average deviation of any agent is 2. Assume that each ST has a reward  $S = 2$ , then the average reward per time-step of all agents is  $\frac{2}{3.2} = 0.625$ . The expected surplus reward of adopting a  $CM = (3, 0.9)$  is given by:

$$\text{ave\_surplus}(3, 2) = 0.9 \times 6 - 0.625 \times 5.6 - \frac{0.625 \times 2}{0.9} = 0.511$$

### 3.3 Deciding how much to bid

When agents receive a request to coordinate they submit a bid based on the amount of reward that they would require to compensate them for deviating from their ST. They also submit their current distance away from the CT square<sup>5</sup>. The agents' bids are also affected by their WtC factor; taking this into account, their bids are based on maintaining the amount of reward they would normally expect to receive factored by  $\omega$ . Other influences on their required reward are the amount of time spent in deviating to the CT square, their average reward per time-step and the probability of success of the CM being proposed.

To formalise this, consider an agent,  $A_i$ , with  $\omega_i$  and average reward per time-step  $r_i$ . The agent calculates its *deviation*, i.e., the number of extra time-steps it requires to reach its ST if it goes via the CT square. Note that if, for example, the CT square lies directly on a path to the ST, the agent's deviation would be zero. Clearly, such an agent will submit a very attractive bid, since the cost to coordinate is effectively zero.

Again by means of illustration consider the agents depicted in figure 1.  $A_1$  at  $[5, 3]$  would take 4 time-steps to reach  $ST_1$  at  $[2, 4]$  directly, but 6 steps going via the CT at  $[3, 2]$ , a deviation of 2 time-steps. However,  $A_2$  at  $[1, 1]$  would take 7 time-steps to reach  $ST_2$  at  $[4, 5]$  directly, and also 7 steps going via the CT at  $[3, 2]$ ;  $A_2$  therefore has a deviation of 0.

To compute the reward  $A_i$  requires from engaging in coordination over the CT, it must be compensated both for its deviation and for the possibility that the CM might fail; it also takes into account its willingness to cooperate:

$$\text{bid}_{ij} = \frac{r_i \times \omega_i \times \text{deviation}_i}{p_j}$$

This formula illustrates how the WtC factor influences an agent. If  $\omega > 1$  it can be described as *greedy*, asking for more reward than it would normally expect to receive; if  $\omega < 1$  it can be described as *self-less*, asking for less than it would normally expect to receive; and, if  $\omega = 1$  it is *neutral*, asking only to be compensated for what it normally expects to receive.

The agent submits its bid to coordinate and its distance from the CT square. If an agent is selected to coordinate, it adopts the CT as its current goal. Its ST is only readopted after the CT has been accomplished; in particular, if it should arrive at the ST square en route for the CT, it does not receive its reward until it returns there.

### 3.4 Deciding which bid to accept

Once the AiC has received bids from all agents it selects that one which maximises its surplus reward, given the new (firm) information it has received. For each agent,  $A_i$ , the AiC knows the time it will take to arrive ( $T_i$ ) and the amount of reward it will require ( $\text{bid}_{ij}$ ). The surplus reward is given by:

$$\text{surplus}_{ij} = p_j \times R - \text{bid}_{ij} - r \times (t_j + T_i)$$

Now, it may be the case that no bids are received which give a positive surplus. Even though the chosen CM had an expected surplus, by chance it may be that no agents are sufficiently near to provide reasonable bids. In such a situation the AiC returns to step [2] of the protocol (figure 2) since, although it has been unlucky, its state is essentially unchanged and attempting to coordinate again is still likely to produce a surplus.

<sup>5</sup> In reality, agents could lie about both of these values. However, such strategic behaviour would not affect the basic decision making processes as they are described here. Thus, at this time, we assume agents bid truthfully.

## 4 EXPERIMENTAL EVALUATION

This section provides an empirical evaluation of the key facets of our decision making framework. The following parameters were fixed for all the experiments: duration (10,000 time units), number of CTs in the grid at any one time (1), number of agents (5) and ST reward (1). The experimental variables were the size of the grid, the reward for CTs, the willingness to cooperate factor and the CMs themselves. Results are shown by averaging figures over 10 runs.

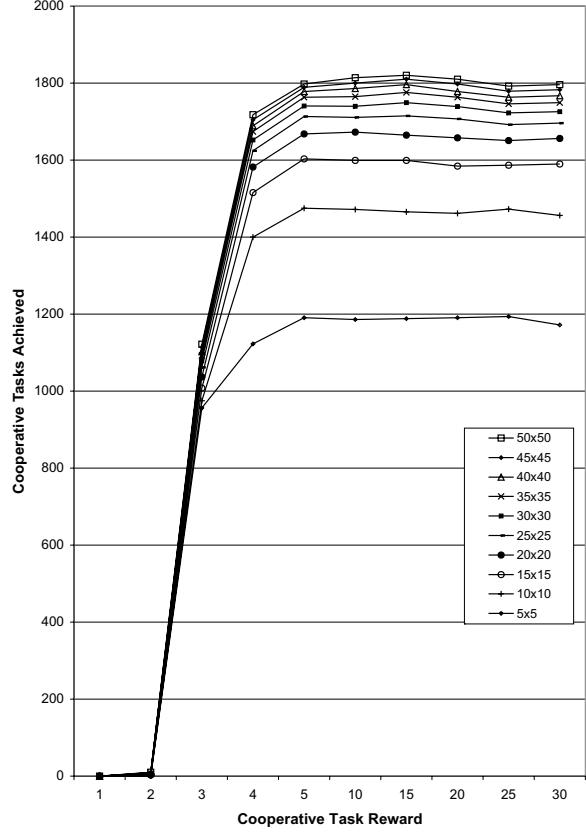
$y \downarrow$	1	2	3	4	5	6	7	8	9	10
$x \rightarrow$	1	2	3	4	5	6	7	8	9	10
1	5	5	5	5	5	5	5	5	5	5
2	5	5	5	5	5	5	4	4	4	4
3	5	5	5	5	4	4	4	4	4	4
4	5	5	5	4	4	4	3	3	3	3
5	5	5	4	4	4	3	3	3	3	3
6	5	5	4	4	3	3	3	2	2	2
7	5	4	4	3	3	3	2	2	2	2
8	5	4	4	3	3	2	2	2	2	2
9	5	4	4	3	3	2	2	2	1	1
10	5	4	4	3	3	2	2	2	1	1

**Figure 3.** Terrain map showing where the CMs are selected.

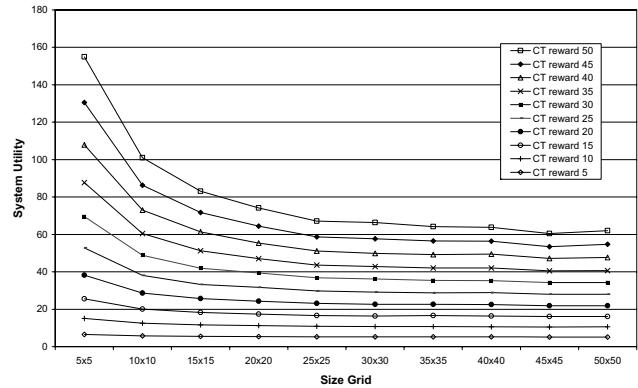
The first experiments sought to show that the agents would indeed choose different coordination mechanisms in different situations. Figure 3 shows the results of which CMs were selected in which grid position. The grid size was 20x20 (the remaining three quadrants are simply a mirror of the portion that is shown and so they are omitted for reasons of space), the reward for CTs is 10, and the agents had the following CMs at their disposal: CM1(0, 0.6), CM2(15, 0.7), CM3(30, 0.8), CM4 (45, 0.9) and CM5 (60, 1.0). In the centre of the grid [10, 10], the agents choose CMs that minimise the set up cost (even though they have a significant chance of failing to ensure coordination). However, as the agents move further away from the centre, so they increasingly prefer mechanisms that are more likely to succeed (even though they have a correspondingly higher set up cost). The explanation for this behaviour is that as the distance from the centre increases, so the expected time for another agent to reach the CT square increases. Thus, to justify its choice of a CT over its ST, the AiC needs to ensure that the cooperations it does enter into do succeed. Whereas, towards the centre of the grid, the time the AiC typically has to wait for another agent to arrive is much smaller and so it can afford to have more cooperations fail. In-between are the points where success and set up time are traded off.

The second set of experiments show that the key determining factor of the amount of cooperation that occurs is the size of the grid. Figure 4 shows that once the reward for CTs is sufficiently high (above 5 in this case) then CTs always get initiated. Hence increasing the reward beyond this figure has no effect. Similarly, figure 5 shows that the system utility (sum of all the agents' utilities) decreases as the grid size increases. Again this is simply because the agents have less opportunity to engage in CTs.

Finally, figure 6 shows the effect of the WtC factor (the reward for CTs is again taken to be 10). As expected, the more greedy the agents become (increasing  $\omega$ ), the fewer the number of CTs that get initiated and achieved. This effect becomes more marked for the CMs



**Figure 4.** Number of Cooperative Tasks achieved using CM(0,1.0).



**Figure 5.** System Utility obtained using CM(0,1.0).

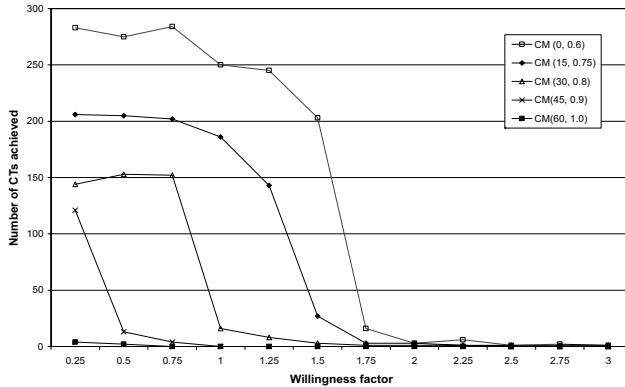


Figure 6. Willingness to Cooperate.

that take longer to set up (because the high set up cost acts as a disincentive to initiate CTs).

## 5 RELATED WORK

The majority of previous work on coordination in multi-agent systems has assumed that it is a design time problem (e.g., [8, 9, 3]). Thus, the designer analyses the system's coordination needs, selects a single coordination mechanism to satisfy these needs, and then imposes this choice upon the individual agents and the overall system. This modus operandi also lies at the heart of mechanism design for multi-agent systems [7]. Here the system designer develops the interaction mechanism (protocol), in order to fulfill particular design objectives, and the agents' decision making behaviours are optimised with respect to the imposed protocol.

The downside of such approaches are that they do not enable an agent to tailor its coordination requirements to its prevailing circumstances. However, as agents are increasingly being used in open and dynamic environments, greater flexibility is needed. To this end, Durfee [2] has argued that agents need the flexibility to coordinate at different levels of abstraction, depending upon their particular needs at a given moment in time. To date, however, this work has focused on building such flexibility into the basic planning mechanisms of the individual agents. As yet, there are no structures or mechanisms for explicitly reasoning about which level to coordinate at in a given situation. Such flexibility was also built into cooperative problem solving agents by Jennings [6]. Here, agents could choose to cooperate according to various conventions which dictated how they should behave in a particular team problem solving context. These conventions varied in terms of the time they took to establish and the communication overhead they imposed upon the agents. However, again, there was no reasoning mechanism for determining which convention was appropriate for a given situation. Boutilier [1] presents a decision making framework, based on multi-agent Markov decision processes, that does reason about the state of a coordination mechanism. However, his work is concerned with optimal reasoning within the context of a given coordination mechanism, rather than actually reasoning about which mechanism to employ in a particular situation.

## 6 CONCLUSIONS

This paper has argued that autonomous agents need to be given the flexibility to dynamically select the mechanism they use for coordinating their actions during collaborative problem solving. Moreover, a formal framework for representing such reasoning has been developed and illustrated in a simple grid-world scenario. Although the specifics of the decision procedures are clearly related to our particular grid-world scenario, we believe that the basic processes and structures we have developed are suitable for reasoning about coordination mechanisms in more general domains. Moreover, the key aspects of our model have been evaluated empirically in order to show their effect on both the individual agents and on the overall system.

For the future, we aim to generalise and extend the reasoning the agents perform about their choice of coordination mechanism. In particular, we will introduce greater heterogeneity both in the agent population and in the coordination mechanisms available, examine the effects of different types of commitment to the cooperative tasks, and use reinforcement learning to enable agents to adapt to their current environment as a substitute for the common knowledge assumption.

## ACKNOWLEDGEMENTS

The second author acknowledges the funding of Mexico's National Council of Science and Technology CONACyT.

## REFERENCES

- [1] C. Boutilier, 'Sequential optimality and coordination in multiagent systems', in *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99)*, pp. 478–485, Menlo Park, CA, (1999).
- [2] E. H. Durfee, 'Practically coordinating', *AI Magazine*, **20**(1), 99–116, (1999).
- [3] E. H. Durfee and V. R. Lesser, 'Partial global planning: A coordination framework for distributed hypothesis formation', *IEEE Transactions on Systems, Man, and Cybernetics*, **21**(5), 1167–1183, (1991).
- [4] J. Filar, *Competitive Markov decision processes*, Springer, New York, NY, 1997.
- [5] J. Galbraith, *Designing Complex Organizations*, Addison-Wesley, 1973.
- [6] N. R. Jennings, 'Commitments and conventions: The foundation of coordination in multi-agent systems', *The Knowledge Engineering Review*, **8**(3), 223–250, (1993).
- [7] J. S. Rosenschein and G. Zlotkin, *Rules of Encounter: Designing Conventions for Automated Negotiation among Computers*, The MIT Press: Cambridge, MA, 1994.
- [8] Y. Shoham and M. Tennenholtz, 'On the synthesis of useful social laws for artificial agent societies', in *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI-92)*, pp. 276–281, San Diego, CA, (1992).
- [9] R. G. Smith and R. Davis, 'Frameworks for cooperation in distributed problem solving', *IEEE Transactions on Systems, Man, and Cybernetics*, **11**(1), 61–70, (1980).