Reasoning about Commitments and Penalties for Coordination between Autonomous Agents

Cora B. Excelente-Toledo
University of Southampton
Dept. of Electronics and
Computer Science
Southampton SO17 1BJ, UK
cbet99r@ecs.soton.ac.uk

Rachel A. Bourne
Queen Mary
Univ. of London
Dept. of Elec. Eng.
London E1 4NS, UK
r.a.bourne@elec.qmw.ac.uk

Nicholas R. Jennings
University of Southampton
Dept. of Electronics and
Computer Science
Southampton SO17 1BJ, UK
nrj@ecs.soton.ac.uk

ABSTRACT

This paper develops and evaluates a new decision theoretic framework in which autonomous agents can make rational choices about coordinating their actions. The framework covers the decisions that are involved in determining when and how to coordinate, when to respond to requests for coordination and when it is profitable to drop contracts in order to exploit better opportunities. Our motivating hypothesis is that enabling agents to dynamically set and re-assess both their degree of commitment to one another and the sanctions for decommitment according to their prevailing circumstances will make coordination more effective. This hypothesis is evaluated, empirically, in a grid-world scenario, taking into account three levels of commitments (total, partial and loose) and three kinds of sanctions (fixed, partially sanctioned and sunk cost).

1. INTRODUCTION

Autonomous agents are increasingly being required to act and interact in dynamic and unpredictable environments. While considerable effort has been devoted to designing agents that can act appropriately in such circumstances, comparatively less effort has been devoted to endowing agents with the ability to *interact* in similarly flexible ways. Thus, for instance, the majority of extant systems impose a single, predefined mechanism that the agents must use for coordinating their actions and agents are often expected to make commitments within such mechanisms that cannot be dropped even if better opportunities present themselves. However, we believe this rigidity is both harmful and unsuitable for many of the applications that are most amenable to an agent-based solution. Thus, this paper reports on our work on endowing agents with decision making procedures that will enable them to coordinate in more flexible ways.

Flexible coordination relies on two key decision making components [7]: (i) the ability to select a means of coordinating that is appropriate to the prevailing situation; and (ii) the ability to assess (and re-assess) an agent's commitment to the on-going coordination activity. Each of these issues will now be dealt with in turn.

With regard to the former point, a variety of protocols and structures have been developed to address the coordination problem. These range from long-term social laws [14], through medium term mechanisms such as Partial Global Planning [6], to one-shot (short-term) mechanisms like the Contract-Net Protocol [15]. All of these *coordination mechanisms* have different properties and characteristics and are suited to different tasks and environments. However, to date, the choice of which mechanism to use is something that the designer imposes upon the system at design time. This means that in many cases the coordination mechanism that is employed is not ideally suited to the agents' prevailing circumstances. To circumvent this problem, we developed a decision making framework that enables agents to dynamically select the coordination mechanism that is most appropriate to their circumstances [2]. While this framework improved both the individual and the system performance, it imposed an overly rigid view of commitment (once agents agree to participate in a coordination activity they cannot renege until it is complete). Therefore the work reported here seeks to extend the preliminary framework by introducing greater flexibility into the handling of commitment and decommitment.

The notion of commitment is central to all coordination mechanisms [7]; it provides the basis of trust on which agents perform their part of the social activity. However the imposition of unbreakable commitments can lead to irrational and inefficient behaviour. For this reason, a number of models have been developed that allow commitments to be dropped if specified contingencies arise (e.g., [4, 7, 9]). While this certainly represents an improvement, the drawback is that the specific conditions under which commitments can be broken must be enumerated in advance. In dynamic and unpredictable environments this can be extremely difficult (and sometimes impossible). To provide yet greater flexibility, *levelled commitment contracts* were introduced [11]. In such contracts, either party can decommit, for whatever reason, as long as they pay the fixed decommitment penalty that is specified in the contract. This type of commitment avoids the problem of having to a priori enumerate specific environmental or agent states and allows agents to decommit unilaterally for whatever reason they deem appropriate. However,

levelled commitments do not explicitly take the ongoing cost of participating in the coordination activity into account. This is because the decommitment penalty is assumed to be fixed, both for the contractor and contractee, no matter at what stage of the coordination process the commitment is broken.

Drawing these points together, it can be seen that flexible coordination requires an agent to make decisions about when to coordinate, which coordination mechanism to use, what levels of decommitment penalty to set, and when to break contracts to take up more promising opportunities. To this end, we present a decision theoretic model that covers all these aspects within the context of a specific coordination scenario based on a grid-world. Such an idealised scenario was adopted to distill an agent's essential decision making capabilities. Nevertheless the scenario incorporates the necessary degree of dynamism and uncertainty to make the models we develop appropriate for more realistic applications. Moreover, the specific models we develop inform the general debate on flexible coordination in that they identify the key reasoning components, specify their inter-relationships, and identify their information requirements.

This work advances the state of the art in two important ways. Firstly, it presents a model that covers the spectrum of reasoning an agent needs to perform in order to coordinate in a flexible manner. No other decision making model is as complete. Specifically, it extends the work presented in [2] in that it incorporates the ability to reason about commitment levels and sanctions for decommitment. Secondly, our framework introduces the notion of *variable penalty contracts*, as an extension to levelled commitment contracts, that incorporate the ongoing cost of participating in the coordination process in the decommitment penalty. Moreover, we present the decision procedures that deal with setting and decommitting from such contracts.

The remainder of this paper is structured in the following manner. Section 2 details our specific coordination scenario. Section 3 describes the model of commitment and penalty used in the scenario. Section 4 formalises the decision procedures of the agents. Section 5 reports on the experimental work to evaluate the effect of the decision making framework. Section 6 deals with related work and Section 7 concludes and presents the areas of further work.

2. THE 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 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 co-ordinating. 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

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

Figure 1: A typical coordination world grid.

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[1] 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 param-eterised 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. If an agent's bid is successful, then it is termed Agent-in-Cooperation (AiCoop) to denote the fact that it is a participant (not AiC) for a CT task. The role Agent-in-ST (AiS) is used to denote the situation where an agent is working toward a ST. The protocol that each agent follows is summarised in figure 2.

In contrast to [2], this scenario permits more than one CT to appear in the grid at any one time and the possibility of more than two agents being required to achieve a CT (this number, m, is specified randomly when the CT is generated). If an agent finds more than one CT in a given cell, it randomly selects one to analyse. Agents might also receive more than one proposal at the same time step, in which case they reply with as many bids as proposals they receive. However, they will only accept one CT contract at a time. Agreements between AiCs and AiCoops to achieve a particular CT are established via a contracting protocol. This contract-net-like protocol consists of three steps. In the first step, AiC broadcasts a proposal to all agents. It then waits for the bids. The second step involves selecting the bids and contracts from AiCs and AiCoops respectively (both of them have to consider refusals and denials of their corresponding offers). Finally, the third step consists of the commitment about the terms of the contract and the time step at which AiCoops will arrive at the CT square.

---

[1]If several agents arrive at a CT square at the same time, one of them is arbitrarily deemed to be in charge.

[1] Agents arrive at a square. If AiS arrives at its ST cell, its goal is attained, it receives the reward and updates its goal. If AiCoop arrives at the CT cell, it notifies the AiC that it has arrived. It might have to wait in the cell until the remaining AiCoops arrive. If AiC receives confirmations from all AiCoops, the CT is achieved and the rewards are paid to AiCoops.

[2] If AiS finds a CT it must decide if it wants to become AiC and, if so, which CM = (t,p) it should use. If t > 0 it must wait t time-steps before broadcasting a request for coordination. If AiCoop finds a new CT, it must decide if it should become AiC or continue with its present aim. If AiC finds a new CT, it ignores it.

[3] If AiS or AiCoop receive a request for coordination, they decide whether and what to bid to participate in the CT. If AiCoop decides to submit a bid, it factors in the penalty fee (if present) for dropping its current contract. The AiC then evaluates all bids. If AiS's bid is accepted, it adopts CT as its new goal. If AiCoop's bid is accepted, it drops its current contract (paying the associated penalty to the AiC) and becomes AiCoop of the new CT. AiC does not respond to requests for coordination.

[4] If AiC receives a decommitment message, it tries to find a replacement for the reneging agent by re-proposing the CT. If it does not receive appropriate bids, it cancels the current CT by paying the contracted penalty to the remaining AiCoops.

[5] Each agent decides on its next move according to its current goal and all agents move simultaneously.

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

As more than two agents may be required to achieve a CT, it is necessary to deal with the fact that an AiCoop may have to wait in the CT cell while the remaining AiCoops arrive (because agents have to travel different distances). In such cases, the AiC pays an additional reward for the time elapsed—AiC knows the number of time steps that each AiCoop is likely to have to wait (specified in the bid) and the amount it will pay for waiting time at a specific waiting rate (q). Thus when an AiCoop notifies the AiC of its arrival at the CT cell, it either receives its share of the CT reward or the waiting rate followed by its share of the CT reward.

3. COMMITMENTS AND PENALTIES

The ability to renege upon commitments and to claim different types of redress impacts the decision making behaviour of both AiCs and AiCoops. In the former case, agents need to be able to attend to and recover from the situation when one of the AiCoops decommits. In the latter case, AiCoops have to assess opportunities to increase their utility by moving to more profitable CTs whenever they are found. To give maximum flexibility in coordination, we believe that agents require the ability to make agreements that involve different levels of commitment and different types of penalty. In particular, we consider three types of commitment:

- Total: Once an agent accepts a contract to achieve a CT, it cannot renege upon it.

- Loose: An agent always drops a commitment if it finds a better option.

- Partial: Agents commit to achieve a CT, but with a percentage of probability they can drop this commit-

ment if they find a better CT to pursue. For example, if an agent has a commitment level of 50%, then half the time it finds a better CT it will cancel and half the time it will continue with its current agreement.

Associated with the dropping of commitment is the use of a penalty model to compensate the agents that remain in the CT. Penalty payments are made each time an agent cancels a commitment and they are paid to the AiC (see section 4.5 for details); here we consider penalties that are:

- Fixed: The amount is fixed at design time and is unrelated to the current coordination context.

- Partially sanctioned: The amount is specified dynamically in the contract when it is agreed. The actual fee depends on the state of the coordination activity and its participants, and the AiC's estimate of the profit that it will receive.

- Sunk cost: The amount is based on the effort that has been invested in the CT to date; if the agents are close to achieving their goal they pay a higher fee than if the contract has only just started.

In identifying whether a new CT is more beneficial than the current one, AiCoops include the decommitment penalty from their existing contract in their deliberations. Thus, a new CT may offer an intrinsically higher reward than the current one, but when the penalty is incorporated the agent may be better off sticking with its existing commitment.

4. AGENT DECISION MAKING

This section formalises the decision procedures of the agents and extends the model in [2] to deal with varying commitment levels and penalties. To study the average impact of coordination mechanisms, an infinite horizon model of decision making was adopted; a finite horizon model may lead to erratic behaviour as the last time-step approaches.

The agents are assumed to be rational and self-interested; so their aim is to maximise their reward, in particular their average reward per unit time. To account for heterogeneity in the population, each agent keeps track of its average reward, termed its *reward rate*, being its total cumulative reward divided by the total number of time-steps taken to obtain it. It uses this rate both to model the (approximate) expected rates of other agents and to decide how much to charge for its own services. Specifically, each agent uses its reward rate to evaluate and compare the different actions available to it; if it can maintain or improve on this rate, it chooses to do so. Of course, this decision model approximates the true relative values of different actions, however, since the environment may change rapidly, we believe that a simple reactive decision procedure is appropriate (see [8]).

There are six types of decisions that agents have to make: the direction to move in; which CM to adopt, if any; how much to bid when a request for coordination is received; how to determine which bid to accept, if any; how much to ask for the penalty fee, and how to determine when to drop a contract.

4.1 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.

4.2 Deciding which CM to select

An agent which, en route to its current goal, encounters a CT, must decide whether to initiate coordination with other agents 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 agent assumes that they are randomly distributed throughout the grid, and that their current goals 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 current goal to reach the CT, while others may be able to coordinate over the CT en route to their own goals. The agent 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 agents (based on the amount of time they must spend deviating from their path and the probability of success of the CM). The expected reward is also affected by the level of commitment in the system (degree_commit); here AiC assumes that its AiCoops have the same degree of commitment as itself. If the agent that finds the CT is currently AiCoop (rather than AiS), then it also needs to subtract the penalty of decommitting from its current contract from the expected reward it expects from the new CT.

The agent 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 rate, r. The agent selects the CM that maximises this surplus reward.

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 takes 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 average distance (ave_dist) away of any randomly situated agent from the CT square, as well as the average deviation (ave_dev) such agents have to make to get there. The agent further assumes that all agents have similar average rewards as its own.[2]

Based on these figures, the agent can assess the expected 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 the other agents to arrive. Since the AiC would usually expect to receive r reward units per time-step, the cost of CM_j is given by:

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

Second, the AiC must estimate the average amount of re-

ward the other m agents will require considering that they have the same degree of commitment as itself:

$$\texttt{ave\_bid}_j(x,y) = \frac{r \times \texttt{ave\_dev}(x,y)}{p_j \times \texttt{degree\_commit}}$$

Third, the agent estimates the expected surplus from adopting the CT taking into account the probability of success of the task, the degree of commitment and discounting the decommitment payment, decommit_cost$_k$ (this includes the current contract value it expects to receive and the penalty it has to pay), from its existing CT$_k$, if there is one.

$$\texttt{surplus}_j(x,y) = p_j \times R \times \texttt{degree\_commit} - \texttt{decommit\_cost}_k$$

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

$$\texttt{ave\_surplus}_j(x,y) = \texttt{surplus}_j(x,y) -$$
$$(\texttt{cost}_j(x,y) + (m \times \texttt{ave\_bid}_j(x,y)))$$

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 the 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 with a total level of commitment. An AiS with total commitment (degree_commit = 1.0), occupies a 5×5 grid and finds a CT requiring one other agent with $R = 6$ at square $[3, 2]$. 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. Given that the agent has an average reward per time-step of $r = 0.625$ and it does not have a penalty to pay, the expected surplus reward of adopting a $CM = (3, 0.9)$ is given by:

$$
\begin{aligned}
\texttt{ave\_surplus}(3,2) &= (0.9 \times 6 \times 1.0) - (3.5 + 1.389) \\
&= 0.511
\end{aligned}
$$

4.3 Deciding what to bid to become an AiCoop

When agents receive a request to participate in a CT they submit a bid based on the amount of reward that they would require to compensate them for deviating from their current goal. They also submit their current distance away from the CT square[3]. To formalise this, consider an agent, A$_i$, with 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 current goal if it goes via the CT square. Note that if, for example, the CT square lies directly on a path to its current goal, 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.

---

[2]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.

[3]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.

To compute the reward A_i requires from engaging in co-ordination over the CT, it must be compensated both for its deviation and for the possibility that the CM might fail or that some of the agents might renege:

$$\texttt{bid}_{ij} = \frac{r_i \times \texttt{deviation}_i}{p_j \times \texttt{degree\_commit}}$$

If the agent is already AiCoop (rather than AiS), $\texttt{bid}_{ij}$ needs to have $\texttt{decommit\_cost}$ added to it.

4.4 Deciding which AiCoop bids to accept

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

Since all AiCoops need to be in the cell at the same time to accomplish the CT, AiC needs to pay an additional award to those AiCoops that have to wait in the CT square. AiC calculates this reward by selecting the agent that will take the longest time to arrive from the set of bids received. From this, it can determine the maximum time, $\texttt{waiting\_time}_{ij}$, that each agent will spend in the cell. For this time, the waiting AiCoops are compensated at a pre-specified rate q. Formally, AiC calculates the $\texttt{cost\_bid}_{ij}$ based on the reward each agent requires and the reward AiC has to pay for the $\texttt{waiting\_time}_{ij}$:

$$\texttt{cost\_bid}_{ij} = \texttt{bid}_{ij} + T_i \times r + \texttt{waiting\_time}_{ij} \times q$$

Next it selects the m bids with the minimum cost:

$$\mathcal{S} = \min[\texttt{cost\_bid}_{ij}]$$

From this $\mathcal{S}$, it selects the furthest bid (i.e., $\max_{i \in \mathcal{S}}[T_i]$) and calculates its expected surplus (which is affected by the likelihood that some of the agents may renege):

$$\texttt{surplus}_j = p_j \times R \times \texttt{degree\_commit} - \sum_{i \in \mathcal{S}}[\texttt{cost\_bid}_{ij}] + r \times (t_j + \max_{i \in \mathcal{S}}[T_i])$$

Now, it may be the case that no bids are received which give a positive surplus. Thus 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 its ST.

4.5 Deciding how to set the penalty fee

The penalty fee can be set independently or dependently of the actual state of the coordination activity. In the former case, the fixed penalty is set as a pre-specified percentage of the CT reward (the actual percentage is an experimental variable that allows penalties to be high, medium or low).

$$\texttt{penalty}_j = \texttt{percentage\_penalty} \times R$$

In the latter case, there are two possibilities. For partially sanctioned penalties, the fee is based on the current expected *surplus* of the CT (again in a proportional manner). In this case, however, the proportion is set according to the degree of commitment within the group; if decommitment is likely, then a high penalty is set since the AiC needs to recoup its costs in setting up and running the group (mutatis mutandis when decommitment is unlikely).

$$\texttt{penalty}_j = (100\% - \texttt{degree\_commit}) \times \texttt{surplus}_j$$

With sunk cost penalties, the fee is calculated with the percentage $\texttt{time\_invested}$ of the CT reward. In contrast to fixed penalties, this percentage is calculated as a ratio of the time spent on the CT and the time that the agent believes needs to be spent in order to complete the CT.

$$\texttt{penalty}_j = \texttt{time\_invested} \times R$$

Partially sanctioned and sunk cost penalties are both variable penalty contracts because the sanction changes dynamically based on the state of the coordination activity.

4.6 Deciding when to drop a commitment

The introduction of loose and partial commitments allows agents to consider the possibility of decommitting. This can occur in two situations: the agent may go from an AiCoop to an AiC or from an AiCoop on its current contract to an AiCoop on a new one (because the $\texttt{bid}_{ij}$ it proposed was accepted). In what follows, let the current CT be subscripted by k and the potential new one by j. Moreover, let $\texttt{expected\_reward}_k$ represent the reward the agent is expecting to get from its current activity (this is calculated as the reward it will receive divided by its distance from the current goal).

An AiCoop will drop its current contract to become an AiC on a new one if it finds a CT such that:

$$\texttt{ave\_surplus}_j(x, y) > \texttt{expected\_reward}_k$$

Similarly, an AiCoop will drop its current contract in favour of becoming the AiCoop on a new CT if:

$$\texttt{bid}_{ij} > \texttt{bid}_{ik}$$

5. EXPERIMENTAL EVALUATION

The experiments seek to explore the following basic hypothesis: incorporating various levels of commitment and penalties into a coordination framework improves the effectiveness of the coordination. The experiments divide into two types: those that probe the effect of the levels of commitment and those that deal with the effect of penalties. To this end, experiments were set-up with a simulation run of 10,000 time steps, a grid size of 10x10, 5 agents in the environment, 3 CTs in the grid at any one time, a maximum group size of 4, a single coordination mechanism, $CM = (1, 1)$, a ST reward of 1 and the reward of the CT was randomly generated in a range of [20, 40, 60]. The experimental variables were the percentage of commitment and the level of penalty which were tested in a range of 0, 25, 50, 75 and 100%. The measured variables were the reward obtained by agent role type, the number of CTs successfully finished, the average penalty fee by penalty type, the number of contracts dropped (CTs decommitted) and how many of them were successfully recovered. The measured variables plotted correspond to the means of the data collected over 10 simulation runs. In all experiments, the means of partial commitment scored significantly higher than those achieved by total or loose commitment[4].

---

[4] Analysis of variance (ANOVA) was used to test hypotheses about differences between the means collected, the null hypothesis of equal means was rejected because the procedure revealed for all experiments that the differences among means were significant, $p < 0.05$ in a range of [.000,.037], $F(2,147)$.
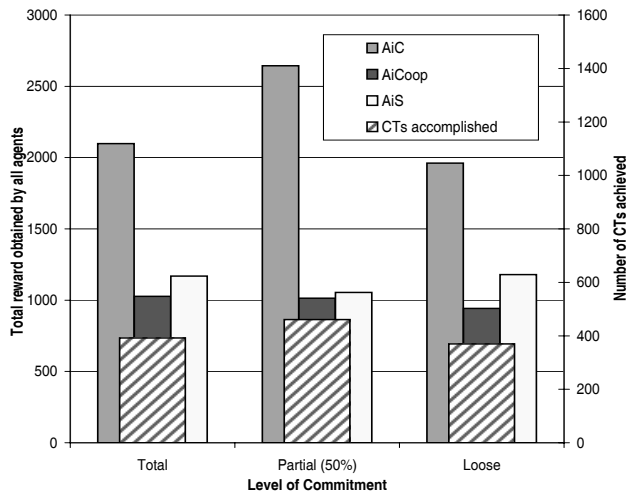
Figure 3: Reward distribution by agent role

Figure 4: Contracts dropped by partial commitment grade

First we deal with the commitment related experiments, here the penalty is set as 25% of CT reward. Figure 3 shows the variance in the reward distribution, by agent role type, for the different levels of commitment. Specifically, the total reward obtained by the agents increases as there are opportunities for decommitment. This is because agents can drop commitments to take up more profitable ones as they arise (even accounting for the fact that they have to pay a penalty). However, looser commitments do not necessarily improve agent performance; if agents can drop commitments easily, then a greater percentage of started CTs fail to finish because agents are continually attracted onto newer more profitable activities. Thus having flexibility to drop commitments is good, but the best performance is achieved by also having some degree of loyalty to existing contracts. The differences between the various levels of commitment in figure 3 are mainly due to the CTs (second axis) accomplished. Here agents with partial commitment get more reward because they achieve more and more profitable CTs; thus AiCs perform better on average. In contrast, AiCoops and AiS receive only the reward they negotiated (which is broadly similar across the commitment types), this is because whether a decommitment occurs or not, they only demand a reward which improves their current situation.

As a consequence of introducing the degree of commitment into the agent's decision making procedures, AiCoops with low levels of commitment have few opportunities to find CTs with high expected surplus rewards (section 4.2) and their bids are too high to ever be contracted (section 4.3). Correspondingly, agents with a high degree of loyalty to CTs have more chances to attempt coordination and their bids have a higher probability of being accepted. Additionally, the degree of commitment affects the frequency with which agents drop contracts, a higher degree means fewer decommitments are performed. Figure 4 clearly illustrates this behaviour. The number of CTs dropped by AiCoops varies with the degree of commitment; AiCoops with a commitment level of 25%, for example, decommit more often than those with 75%. In contrast, AiCoops with a commitment level of 0% even though they have more opportunities to do so, decommit much less frequently because their bids to participate in new CTs are too high (as previously noted).
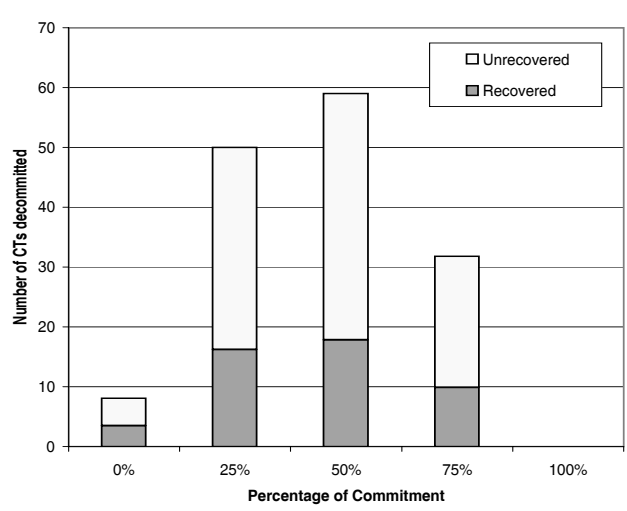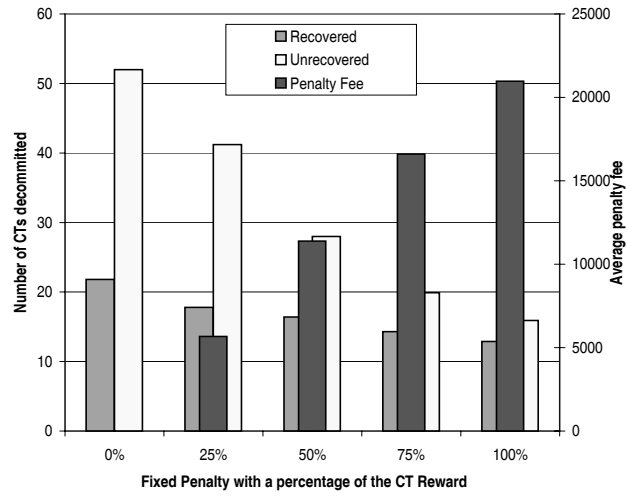

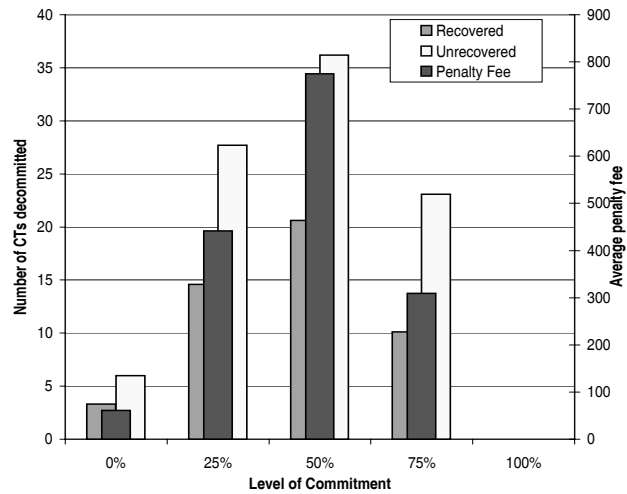**Figure 5: Fixed penalties**

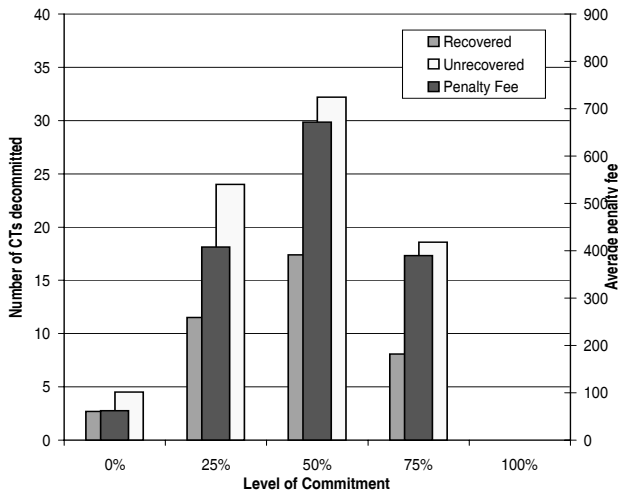**Figure 6: Partially sanctioned penalties**
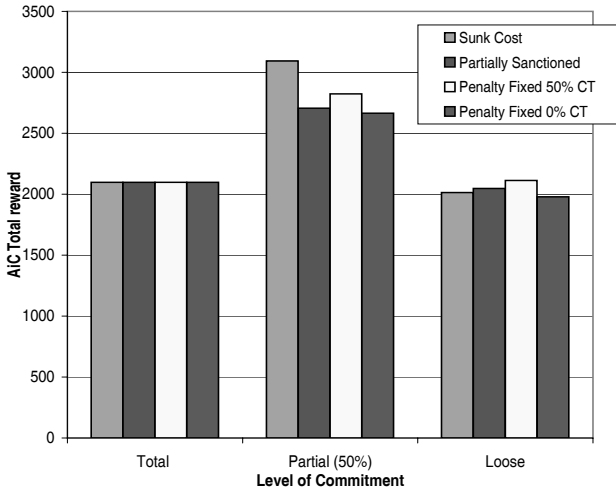
Figure 7: Sunk cost penalties

Figure 8: Comparing fixed, partially sanctioned and sunk cost penalties

Turning to the penalty related experiments. Figures 5, 6 and 7 show the effects of the number of contracts decommitted using fixed, partially sanctioned and sunk cost penalties, respectively. With fixed penalties, figure 5 indicates that the number of contracts dropped by AiCoops (here AiCoops have a commitment level of 50%) gradually decreases as the penalty fee increases (because agents cannot afford a high penalty fee). With partially sanctioned (figure 6) and sunk cost penalties (figure 7), the same broad trends can be observed, fewer decommitments occur with low and high degrees of commitments. This is because of the high penalties for decommitment and as a consequence of the fewer opportunities to decommit which occur when an agent has a low degree of commitment. However, both figures show similar values for the average penalty fee (second Y axis) and for the number of decommitments performed. In short, AiCoops with lower penalties perform more decommitments independently of the kind of sanction that is in place.

To compare the three penalty types, figure 8 shows the AiC total reward by commitment level type. The best reward is obtained by combining a sunk cost penalty with par-

tial commitment (50%). With sunk cost and partially sanctioned penalties, agents decommit less frequently than they do with a 0% fixed penalty but more often than they do with a 50% of penalty fixed, however the reward obtained by AiCs using a fixed penalty of 50% is better than that with partially sanctioned penalties. Using a sanction which changes dynamically with the state of the coordination, specifically, sunk cost, seems a good model because it sets a more appropriate payment for decommit when compared with partially sanctioned penalties. In sum, as was the case with the commitment related experiments, sunk cost and partial commitment level agents represent the most effective combination between loyalty to the existing situation and flexibility to discover better situations and to submit better proposals.

6. RELATED WORK

There are two broad strands of work that are primarily related to our model: (i) work on reasoning about coordination; and (ii) work on commitments to contracts. Each will now be dealt with in turn.

In terms of coordination, most extant work assumes it is a design time problem (e.g., [14, 15, 6, 10]). Thus, comparatively little work addresses run-time reasoning about the selection of particular coordination protocols. Durfee [5] 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, there are no mechanisms for explicitly reasoning about which level to coordinate at in a given situation. Such flexibility was also built into cooperative agents by Jennings [7]. Here, agents could choose to cooperate according to various conventions which dictated how they should behave in particular team problem solving contexts. 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. Barber et al. [1] present a software engineering framework that enables agents to vary their coordination mechanisms according to their prevailing circumstances. They also identify criteria for determining when particular mechanisms are appropriate. However, the decision procedures for actually trading-off these criteria are not well developed. Boutilier [3] 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.

In terms of work on commitment, our model is most closely related to that of Sandholm and Lesser's levelled commitment contracts [11] (a discussion of how our model relates to that of other work on commitment is given in section 1). Our approach builds upon Sandholm's basic intuition that agents should be able to unilaterally decommit from a contract, for whatever reason they deem appropriate, subject to the payment of a penalty. However, our model differs in a number of important ways. Firstly, we cover more than just reasoning about decommitment. Secondly, levelled commitment contracts typically assume a fixed penalty for decommitment that ignores the current costs of the ongoing coordination activity. Our variable penalty contracts thus offer a more realistic model for assessing the real cost of

reneging. Thirdly, Sandholm's original proposal contained no algorithms (decision procedures) for agents to compute when they should decommit from a given contract. This was rectified in [12], however only in a limited manner. In particular, his algorithm for computing the Nash equilibrium decommitment threshold relies on the fact that agents have information about the actual and likely alternative options (as well as their probability distribution functions) that may be presented to the agents with which they are coordinating. This is, we believe, somewhat unrealistic and is not required by our model. Sen and Durfee's work on commitments in the domain of distributed scheduling is also related [13]. They focus on the use of different commitment strategies and evaluate the impact of various environmental factors on their effectiveness. Their results show that the "blind" use of commitment in their system does not improve the performance. This is also consistent with our experimental findings, but they do not have our richness of commitment and penalty types to enable the agents to overcome this shortcoming.

7. CONCLUSIONS

This paper has developed a decision theoretic framework that enables agents to coordinate in a flexible manner. In particular, we have focused on the issues of variable commitment levels between agents and of different penalty sanctions for reneging on contracts. Our empirical results have highlighted the fact that flexibility with respect to commitment levels can improve the effectiveness of coordination. We showed that a certain degree of loyalty to existing contracts leads to better overall performance than continually jumping to new opportunities as they arise. For penalty sanctions, we showed that setting them based on the prevailing context also improves coordination.

For the future, we aim to incorporate our empirical findings into the agent's decision procedures so they can select the level of commitment and penalty sanction for themselves according to their prevailing circumstances. To account for more heterogeneous agent populations, we also intend to allow agents to adapt their decision making to reflect their experiences of the individual agents in the system. This will enable agents to make more realistic predictions about the aims and behaviour of their potential collaborators.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] K. S. Barber, D. C. Han, and T. H. Liu. Coordinating distributed decision making using reusable interaction specifications. In *Third Pacific Rim International Workshop on Multi-Agents*, pages 1–15, Melbourne, Australia, 2000.

[2] R. A. Bourne, C. B. Excelente-Toledo, and N. R. Jennings. Run-time selection of coordination mechanisms in multi-agent systems. In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI-2000)*, pages 348–352, August 2000.

[3] C. Boutilier. Sequential optimality and coordination in multiagent systems. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99)*, pages 478–485, Stockholm, Sweden, 1999.

[4] P. Cohen and H. J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.

[5] E. H. Durfee. Practically coordinating. *AI Magazine*, 20(1):99–116, 1999.

[6] 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.

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

[8] D. N. Kinny and M. P. Georgeff. Commitments and effectiveness of situated agents. In *Proceedings of the Twelfth InInternational Joint Conference on Artificial Intelligence (IJCAI-91)*, pages 82–88, 1991.

[9] H. Raiffa. *The art and science of negotiation*. Harvard University Press, 1997.

[10] J. S. Rosenschein and G. Zlotkin. *Rules of Encounter*. The MIT Press: Cambridge, MA, 1994.

[11] T. W. Sandholm and V. R. Lesser. Advantages of a leveled commitment contracting protocol. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI-96)*, pages 126–133, Portland, Oregon, August 1996.

[12] T. W. Sandholm, S. Sikka, and S. Norden. Algorithms for optimizing leveled commitment contracts. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99)*, pages 535–540, Stockholm, Sweden, 1999.

[13] S. Sen and E. H. Durfee. The role of commitment in cooperative negotiation. *International Journal on Intelligent & Cooperative Information Systems*, 3(1):67–81, 1994.

[14] 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)*, pages 276–281, San Jose, California, July 1992.

[15] 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.