Norm Adaptation and Revision in a Multi-Agent System

Nick Lacey and Henry Hexmoor
University of Arkansas
313 Engineering Hall
Fayetteville, Arkansas
{nlacey,hexmoor}@uark.edu

Abstract

In this paper we address the question of assigning social norms to agents: should we attempt to ascribe social norms to agents that will act in complex dynamic environments, or is it possible to allow the agents to adapt to new situations as they arise, and choose their norms accordingly? We argue that adaptation is preferable to prescription, in that agents should be allowed to revise their norms on the fly. A system is constructed in which the performance of multiple agents operating in the same environment can be assessed. Experimental results concerning alternative norm selection strategies are presented and discussed.

Introduction

As agent designers, we want our agents to be able to operate reliably in dynamic environments. The environment may change in ways which are not only more complex than the agents are able to model, but also too complex to allow us to fully predict how an agent's actions will interact with its environment. This means that we cannot simply hard-wire preprogrammed responses to certain situations into our agents: rather, they have to be *responsive*.

This is especially true in multi-agent systems: the complexity involved in attempting to foresee every possible way that a single agent could interact with its environment clearly increases as more agents are added to the system.

In this paper we address the question of assigning social norms to agents: should we attempt to ascribe social norms to agents that will act in complex dynamic environments, or is it possible to allow the agents to adapt to new situations as they arise, and choose their norms accordingly? We argue that flexible adaptation is preferable to prescription: rather than attempting to guess the norms that will be required at run time, agent designers should allow their agents to dynamically revise their norms in response to changes in the environment.

This paper is structured as follows: We begin by describing the motivation for this research. We then describe the implementation that was carried out, as well as the results that were achieved. Finally, we offer some conclusions that can be drawn from this research, and suggest directions for further research.

Representing and Regulating Social Behaviour Using Norms

Norms are conventions which facilitate the co-existence of multiple agents within the same environment. For example, when driving in the USA, it is useful to uphold the norm of driving on the right hand side of the road. In the UK, on the other hand, the adoption of the norm that one should drive on the left hand side of the road is likely to prove more beneficial.

The above example is useful as it shows that, while norms can be essential to the successful completion of agents operating within a social environment, they can also be completely *arbitrary*. Some countries drive on the left, while others drive on the right; most drivers neither know nor care *why* their country has adopted the norm it has chosen, and as long as all drivers uphold the norm, it does not matter that the individual agents do not know its origins.

The arbitrary nature of norms can be something of a double-edged sword for agent designers. On the one hand, norms are presented as cognitive modalities which must be upheld, rather than justified. This can make deep modelling and norm recognition difficult. For example, if agent A perceives agent B driving on the left hand side of the road, further contextual information is required before agent A can determine whether or not agent B is upholding or ignoring the relevant norms.

In this example, human agents are usually aware of which country they are driving in, and are able to infer whether or not other drivers are upholding norms relatively easily. However, even reasoning that is trivial for humans can be expensive for artificial agents, both in terms of execution speed and in terms of the extra background knowledge required.

On the other hand, the fact that rational agents can follow norms without having to consider *why* they are following them can be very useful to agent designers, as it greatly facilitates the process of *norm revision*. This is because the agent designer is able to program her agents so as to benefit from the social advantages of following norms without having to worry about whether or not the adoption of a norm is consistent with the agent's existing beliefs.

Allowing agents to revise their norms on the fly allows greater flexibility than prescribing which agents should follow which norms at which time in advance, as it allows the agents, and hence the society as a whole, to react to changes in the environment. Except in trivial cases, the Frame Problem prevents agent designers from being able to predict *exactly* how a complex environment will change over time.

Norm Revision and the Prisoner's Dilemma

The Prisoner's Dilemma is a zero-sum game, in that the aim of the players is not to minimize the utility of their opponent, but to maximise their own utility. This means that it is possible for both players to achieve the same utility, or for either player to achieve higher utility than the other.

		Player 2	
		Cooperate	**Defect**
Player 1	**Cooperate**	3,3	0,5
	Defect	5,0	1,1

Table 1: Utilities in the Prisoner's Dilemma

Table 1 shows the payoffs associated with the Prisoner's Dilemma (Sandholm 1999). In this case the payoffs represent arbitrary positive units, unlike in some versions of the Prisoner's Dilemma, where the payoffs represent arbitrary negative units. This table shows that if both players cooperate, both players receive a utility of 3, while if Player 1 cooperates and Player 2 defects, Player 1 receives a utility of 0 while Player 2 receives a utility of 5.

In the traditional Prisoner's Dilemma, players play each other once only. In the Iterated Prisoner's Dilemma, the same players play each other repeatedly. An interesting factor in the Iterated Prisoner's Dilemma is that if a player constantly chooses the same option, her opponent may be able to predict what she is about to do and exploit the situation. For example, if Player 1 constantly cooperates, and Player 2 realises this, then Player 2's best option is to constantly defect, as this will ensure maximum utility for Player 2.

For this reason, in the Iterated Prisoner's Dilemma players are not so concerned about choosing the best move in any particular turn, but are concerned about choosing the best *strategy* in order to respond to their opponent's choices. We therefore now turn our attention to the representation and manipulation of social strategies in rational agents.

In (Dignum *et al.* 2000), Dignum *et al.* argue that integrating social norms into the standard BDI approach can yield socially aware rational agents. According to the BDI approach to agent design (Wooldridge 2000), the cognitive functions of a rational agent are categorised into the following three modalities:

Beliefs are facts which the agent holds which represent the properties about the agent's environment. Ideally, the agent's current belief set should be consistent.

Desires are the agent's long term goals. There is no requirement that the agent's desires should be consistent.

Intentions represent a staging post between beliefs and desires, in that they represent goals or sub-goals that the agent intends to actually bring about.

Dignum *et al.* argue that norms are the "glue" that bind autonomous agents together in a multi-agent system.[1] They also say that, as circumstances within the environment and other agents in the system may change, these norms should not be hardwired but should be flexible.

Elsewhere we put forward an approach to belief revision in multi-agent environments based on social notions, intentions, and constraint satisfaction (Lacey, Hexmoor, & Beavers 2002) (Lacey & Hexmoor 2002). Our approach is based on the *coherence* approach to belief revision (Doyle 1992). Thagard and Millgram (Thagard & Millgram 1997) suggest an interesting approach to decision making based on coherence. They define decision making as inference to the best plan, and suggest that when people make decisions they do so by adopting *"... complex plans on the basis of a holistic assessment of various competing actions and goals"*

Like Thagard and Millgram, we suggest that decision making involves complex planning based on the assessment of competing holistic interpretations. However, whereas Thagard and Millgram suggest a connectionist approach to modeling this process, we have used a symbolic approach which allows us to explicitly model and represent different classes of belief, such as actions, perceptions, explanations, and intentions.

Implementation and Experiments

A system in which multiple BDI agents interact was implemented in SICStus Prolog (Carlsson & Widén 1993). The graphical output reproduced in Figure 1 was produced using the SICStus Tcl/Tk library (Almgren *et al.* 1993).

The experimental domain that was chosen was based on the concept of agents delivering goods in a city. The number of completed journeys made by each agent is recorded, as is the total number of journeys made by all the agents. The efficiency of an individual agent is measured in terms of the number of journeys it is able to complete, while the efficiency of the entire system is measured in terms of the total number of journeys that are completed.

As shown in Figure 1, the environment is composed of 8 terminals, labeled as `t1` ... `t8`, and 4 junctions, labeled `j1` ... `j4`. At each junction an East-West road crosses a North-South road. The junction is controlled by a traffic signal which allows traffic to flow either East-West or North-South. The bars at each junction show the direction that traffic is allowed to flow. For example, when the environment is in the state shown in Figure 1, traffic is flowing North-South and `j1` and `j2`, and East-West at `j3` and `j4`.

Agents start out at randomly assigned terminals, and must make their way to another randomly assigned terminal. Depth-first search is used to find a suitable route from one terminal to another.

When an agent reaches a junction which is blocked to traffic coming from its direction, the correct course of action is of course for the agent to stop and wait for the signal to change. Whether or not the agent acts in this ways depends

[1]Dignum *et al.* actually consider both norms and obligations. For the sake of simplicity we have considered obligations to be implicit as far as this paper is concerned.

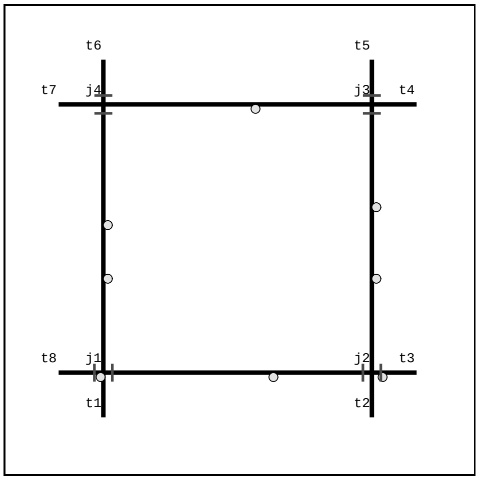

Figure 1: The Experimental Domain

upon the norm that it is following. There are two norms that agents may adopt, which we have labeled *cooperate* and *defect*:

Cooperate The agent obeys all traffic signals. The journey of a cooperating agent will usually take longer than that of a defecting agent, as cooperating agents have to wait for the signals to change before proceeding.

Defect The agent completely disregards all traffic signals. Defecting agents usually complete journeys faster than cooperating agents. However, defecting agents can cause consequences that are detrimental to the community as a whole.

As defecting agents ignore traffic signals, there is a chance that a defecting agent will go through a stop signal, and will collide with another agent. Collisions have two consequences:

Individual Penalty The agent that *caused* the collision, that is, the agent that disobeyed the stop signal, remains stopped for the duration of the penalty. The innocent party involved in the collision is not affected by this penalty.

Group Penalty The entire junction where the collision occurred is closed for the duration of the penalty.

Parameter	**Value**
Number of Agents	50
Number of Timeslices	1000
Group Penalty	25
Individual Penalty	25
Signal on Red(%) (NS,EW)	50,50

Table 2: Summary of Experimental Parameters

The values of these penalties, as well as of the other major parameters used for the experiments described in this paper are summarised in Table 2. The values next to 'Signal on Red' refer to the percentage of timeslices when any given traffic signal will be red for traffic traveling North-South and East-West. The value of (50,50) means that each signal allows traffic to flow North-South for 50 timeslices, and then allows traffic to flow East-West for 50 timeslices. Right turns on red are not permitted.

Experiments

Given the parameters summarised in, Table 2, the obvious question to consider is, out of the 50 agents in the environment, how many should be cooperating?

If too many agents are cooperating, there will be no collisions and hence no penalties, but most of the agents will spend too much time waiting at traffic signals. On the other hand, if too few of the agents are cooperating, there will be many collisions and penalties, resulting in junctions being closed to all agents. In this section we investigate different norm selection strategies, and examine how many journeys are completed under each strategy. The norm selection strategies we examine are:

1. Prescriptive Cooperation
2. Prescriptive Defection
3. Adaptive Convergence
4. Adaptive Divergence

Table 3 shows the utilities that would be exhibited by an agent operating in this environment in various circumstances. A '+' indicates a probable increase in performance, whereas a '-' indicates a probable decrease in performance. A '0' indicates that there will probably be no significant change in performance.

		Group	
		Cooperate	**Defect**
Individual	**Cooperate**	-,0	+,+
	Defect	+,0	-,-

Table 3: Individual and Group Utilities for the Delivery Scenario

If the majority of agents are cooperating, then an individual agent will probably harm its performance by cooperating. This is because the chances of its being involved in a collision are small compared to the significantly decreased journey time. Similarly, the effects on the group's performance of one agent defecting or cooperating will probably not be significant if the majority of agents are cooperating.

On the other hand, if the majority of agents are defecting, then the society as a whole has problems, as there will be a high likelihood of collisions, which will adversely affect global performance. In this situation, both individual and society as a whole will probably benefit from the cooperation of the individual, as this could be the first step in a move from anarchy to order.

The major difference between the situation reflected in Table 3 and the traditional Prisoner's Dilemma summarised in Table 1 is that the Prisoner's Dilemma involves the interaction of one agent with one other agent, while Table 3 shows the utilities that an individual agent would exhibit when interacting with *a community* of agents. The asymmetries in this table stem from the fact that an individual is nonetheless part of the group. As such, if the actions of the individual adversely affect the performance of the group, then its actions will ultimately adversely affect its own performance as well.

Another difference between the Prisoner's Dilemma scenario and Table 3 is that in the Prisoner's Dilemma, the results of an agent's choice of strategy are detectable immediately, while in this situation the effects of an agent moving from a cooperating norm to a defecting norm in the delivery environment may not become apparent for several timeslices.

Despite these differences, both tables represent a zero-sum situation in which the goal of the participating agent is to maximise its utility rather than minimising the utility of the other agents in its environment.

Prescriptive Cooperation Under this strategy, all agents in the system cooperate all the time. This means that all agents will always stop for red lights, and as such there will be no collisions.

Prescriptive Defection Under this strategy, all agents defect all the time. This means that no agents will stop for red lights. As such some collisions are to be expected.

Adaptive Convergence Agents whose performance is lower than average defect, while those whose performance is higher than average cooperate. This strategy was named to reflect the fact that under-performing agents would defect, and hence be able to complete their journeys more quickly. Once their performance had exceeded that of the system average, they would revert to a cooperative norm. This will result in the relative performance of individual agents converging towards the average over time, with little variation in the performance of individual agents.

Adaptive Divergence Under this strategy, agents whose performance is higher than average defect, while those whose performance is lower than average cooperate. This means that agents who are performing well will continue to defect and probably perform well, while those who are performing poorly will continue to cooperate. This will result in the relative performance of individual agents diverging from the average performance over time. This means that we can expect that systems operating on a divergent strategy would show more variation in the performance of individual agents than would be present in systems operating under a convergent strategy.

Results

Table 4 summarises the results that were obtained from these experiments. The "Journeys" column lists the total number of journeys achieved under the various strategies. The third column lists the Standard Deviation (σ) of the number of journeys completed for each agent, while the fourth column lists the number of collisions that were recorded under each strategy.

The first result to note is that the number of completed journeys is higher when the agents are able to revise their norms, as they are under the Convergent and Divergent strategies, than when all agents are either cooperating or defecting.

As would be expected, if all the agents in the system are cooperating, there are no collisions and hence no penalties, but the amount of time that the agents spend waiting at traffic signals results in fewer completed journeys than result from the alternative strategies.

Norm	Journeys	σ	Collisions
Cooperate	360	0.41	0
Defect	372	0.48	100
Convergent	391	0.41	38
Divergent	390	0.54	26

Table 4: Summary of Results

On the other hand, if all the agents are defecting, the agents might be able to finish they journeys more quickly, but there is a much greater chance of collision: 0.29 collisions per completed journey. Also, the fact that the standard deviation is higher for the defect strategy than for the cooperate strategy shows that the variation in the level of performance is higher for the defect strategy.

These results show that the adaptive strategies performed better than the prescriptive strategies. Both the Convergent and Divergent strategies resulted in a higher number of completed journeys that the prescriptive strategies, and fewer collisions than the Defect strategy. The Convergent strategy averaged 0.10 collisions per completed journey, while the Divergent strategy averaged 0.07 collisions per completed journey.

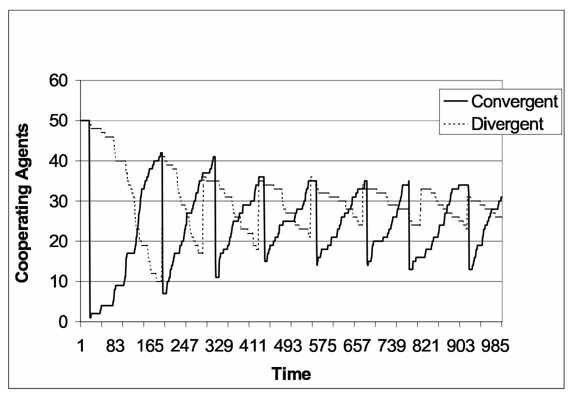

Figure 2: The Number of Cooperating Agents Over Time

Thus, under the adaptive strategies, the agents are able to revise their norms in response to changes in the environment. Figure 2 shows the number of cooperating agents over time for both the Convergent and the Divergent strategies.

This figure shows that under both strategies, there is initially considerable variation in the number of cooperating agents, but this variation gradually subsides over time, leading to a more stable environment. Thus, around the 200th timeslice, the number of cooperating agents goes from 10 to 42 in the Divergent strategy, but by the 900th timeslice, the change is from 25 to 32.

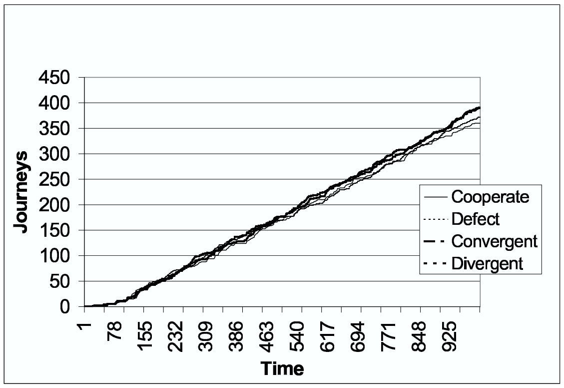

Figure 3: The Total Number of Journeys Completed Under Different Strategies

While the number of Cooperating agents varies over time, the rate of increase in the number of journeys remains relatively constant. Figure 3 shows the total number of journeys over time for all four strategies. While this figure is useful in that it shows that the rate of change is relatively constant over time, we appreciate that due to the similarity in the results for the four strategies, the figure is not particularly informative beyond this. For this reason we have also included Figure 4, which concentrates only on the final 300 timeslices of Figure 3.

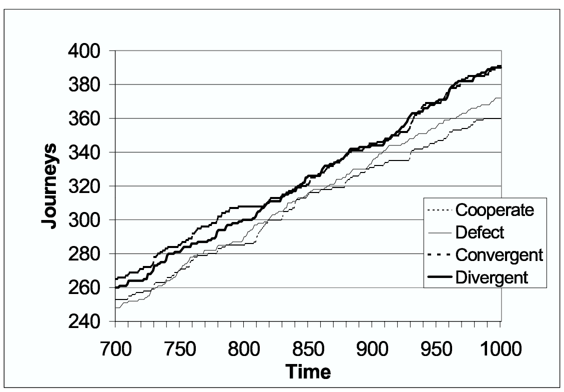

Figure 4: A Summary of Figure 3 Emphasising the Differences in Performance Between the Various Strategies

Figure 3, in conjunction with Table 4, shows that the two adaptive strategies perform better than the two prescriptive strategies.

Conclusions and Further Work

In this paper we have argued that the representation of social norms within the knowledge base of an agent which is to act within a multi-agent system allows the system designer to control the social behaviour of the agents in the system.

We have also argued that in some situations it is more efficient to allow the agents in the system to adapt to new situations by revising their norms as appropriate, than it is to attempt to prescribe norm adherence at design time. We describe the implementation and testing of a system in which multiple BDI agents capable of norm revision interact. Experimental results show that in some circumstances adaptive norm revision strategies perform better than prescriptive norm assignment at design time.

We do not, however, claim that this will always be the case, as there will be situations in which agents should not be able to revise their norms. For example, in safety critical systems, we may wish to be able to predict exactly what an agent will do in a given situation, without giving it the freedom to change its behaviour as it sees fit.

The next stage in this research is therefore to conduct further experiments in an attempt to identify situations in which norms should be prescribed at run time, and when adaptive norm selection should be used.

References

Almgren, J.; Andersson, S.; Carlsson, M.; Flood, L.; Haridi, S.; Frisk, C.; Nilsson, H.; and Sundberg, J. 1993. *SICStus Prolog Library Manual.* Swedish Institute of Computer Science, Kista, Sweden.

Carlsson, M., and Widén, J. 1993. *SICStus Prolog User's Manual*. Swedish Institute of Computer Science, Kista, Sweden.

Dignum, F.; Morley, D.; Sonenberg, E. A.; and Cavedon, L. 2000. Towards socially sophisticated bdi agents. In *Proceedings of the Fourth International Conference on Multi-Agent Systems (ICMAS?2000)*.

Doyle, J. 1992. Reason maintenance and belief revision. Foundations vs Coherence theories. In Gärdenfors, P., ed., *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge: Cambridge University Press. 26–51.

Lacey, N., and Hexmoor, H. 2002. Representing and revising norms and values within an adaptive agent architecture. In *The 2002 International Conference on Artificial Intelligence IC-AI 02*.

Lacey, N.; Hexmoor, H.; and Beavers, G. 2002. Planning at the intention level. In *15th International FLAIRS Conference*. Pensacola Beach, Florida: FLAIRS.

Sandholm, T. W. 1999. Distributed rational decision making. In Weiss, G., ed., *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, MA, USA: The MIT Press. 201–258.

Thagard, P., and Millgram, E. 1997. Inference to the best plan: A coherence theory of decision. In Rom, A., and Leake, D., eds., *Goal-driven learning*. Cambridge, MA: MIT Press. 439–454.

Wooldridge, M. 2000. *Reasoning About Rational Agents*. Cambridge, MA: MIT Press.