Variable Sociability in Agent-Based Decision Making

Lisa Hogg and Nicholas R. Jennings

Department of Electronic Engineering,
Queen Mary and Westfield College
London, England, E1 4NS
{l.m.hogg, n.r.jennings}@elec.qmw.ac.uk

Abstract. Multi-agent system research is concerned with the issues surrounding the performance of collections of interacting agents. A major concern, therefore, is with the design of the decision making mechanism that the individual agents employ in order to determine which actions to take to achieve their goals. An attractive and much sought after property of this mechanism is that it produces decisions that are *rational* from the perspective of the individual agent. However, agents are also inherently social. Moreover, individual and social concerns often conflict, perhaps leading to inefficient performance of the individual and the system. To address these problems we propose a formal decision making framework, based on social welfare functions, that combines social and individual perspectives in a unified and flexible manner. The framework is realised in an exemplar computational setting and an empirical analysis is made of the relative performance of varyingly sociable decision making functions in a range of environments.

1 Introduction

An increasing number of applications are being implemented as multi-agent systems in which individual agents interact with one another in pursuit of both individual and system goals. When building such systems, a number of fundamental design questions continually arise: what underlying motivations should the agents possess to achieve the required overall results?; how should agents make decisions given that they inhabit a multi-agent environment and that their choices effect others?; and to what extent should individual agent autonomy be sacrificed for global concerns? Answers to these basic questions are needed if robust and reliable agent-based systems are to be developed.

Designers want their agents to be rational and to do the right thing [15]. To this end, a major strand of research has adopted the economic viewpoint and looked at self-interested agents [6] that consider what action to take solely in terms of its worth to themselves. However, this is only part of the story. When an agent is situated in a multi-agent context, its actions can have non-local effects. For example, the actions of different agents can conflict or result in duplication of action. This can lead to undesirable results and inefficient utilisation of common resources. Consequently, the benefits, to both the individual and the overall system, of a more social perspective on decision making are beginning to be realised [1, 7] and notions of social rationality [9] are emerging. This notion of socially acceptable decisions has long been of interest in the field of socio-economics [11]. However, to date, there has been little cross fertilization with agent based computing research. One of the main aims of this research is to examine this link and explore the conflict between individual and social concerns in a multi-agent systems.

In addition to balancing individual and social concerns, agents typically need to operate in a resource bounded manner. They do not have unlimited time or computational resources. Moreover, such bounded rationality should be responsive to fluctuations in the amount of resources available. Hence, agents should be able to modify how they make decisions based on their current context. In our case, this means agents should be able to dynamically vary their balance between individual and social considerations depending on the amount of resources available in the system. Moreover, because computing the social effects of action choices consumes resources, agents need to be able to vary the effort they expend on this task. Thus, when resources are plentiful an agent may wish to expend a significant amount of effort computing the social implications of an important choice. However, when resources become scarce, the same agent may choose to adopt a computationally cheaper approach to the same decision.

As a first step towards this long term vision, this paper investigates, in an empirical fashion, the feasibility and performance of socially rational decision making. We define a decision making framework based on work found in socio-economics that explicitly characterises how agents can determine which action to perform in terms of a balance between individual and social concerns. By being explicit about the constituent components, the framework provides the flexibility to enable agents to dynamically tune their operation in order to be as rational as possible in the prevailing circumstances. In particular, this paper extends previous work in two main ways. Firstly, it describes a marriage of socio-economic and agent-based techniques to investigate how social reasoning can be effectively employed by an agent situated in a multi-agent system. Secondly, it stimulates ideas as to how social reasoning agents may be controlled in times of resource constraints, allowing inroads to be made in the construction of socially rational, resource bounded reasoners.

The remainder of the paper is structured as follows. Section 2 details the socially rational decision making framework and introduces the multi-agent platform used in our empirical evaluation. Section 3 describes the experiments we performed to assess our hypotheses about socially rational decision making. Related work is discussed in section 4, followed by the conclusions and future work in section 5.

2 Balancing Conflicting Needs

To date, the dominant decision making philosophy in agent design has been to equate rationality with the notion of an individual maximising a self-biased utility function. Thus, an agent's motivation is the maximisation of benefits with regards to its own goals. However, in multi-agent systems, for the reasons outlined above, a more social perspective on decision making is often desirable. Traditionally, this has been achieved by making the overall system the primary unit of concern. This has the consequence of subordinating an agent's autonomy to the needs of the system. For this reason, we believe such top-down approaches fail to exploit the full potential of the agent-oriented approach; therefore we propose an alternative means of achieving the same end. Thus we wish to build agents from the micro to macro level, but still retain the benefits of a more social perspective. To this end, our approach is to incorporate an element of social consideration into each agent's individual decision making function.

One means of achieving good system performance from the micro level is to incorporate all the necessary social information into a single, amorphous utility function. This is the method that would be followed by advocates of traditional decision theoretic approaches. However, such an approach conceals important details of how (and why) the agent actually reasons. Such details are not only important for the analysis of agent behaviour, but also provide a vital tool to designers when building complex systems. Therefore we advocate an approach that provides detailed guidance as to how social agents may be constructed.

Making the theory of social decision making finer grained in this manner is also essential for progress on the issue of bounded social rationality. Here, parallels can be drawn between conceptions of meta-reasoning [14] and the idea of controlling the amount of social reasoning that should be performed by contracting and expanding the set of acquaintances the agent considers in its reasoning.

2.1 A Social Decision Making Framework

In order to ascertain the social impact of an action, an agent needs to be able to determine the value that a state (as a result of an action) has for other individuals and possibly for the whole society. To do this, the agent needs to empathise with what others value (i.e. know how others value states and be able to make interpersonal comparisons). In this case, the social decision framework developed here builds upon and extends the idea of social rationality proposed by Jennings and Campos [9] and is based on Harsanyi's social welfare function [8]. Social welfare functions were first introduced by sociologists and they deal with choice by a group of individuals in a society. The decision maker can either be a group making a joint decision or an individual making a choice that has global consequences. The general theory of social welfare is formalised as follows. A set of individuals I $\{1, ..., n\}$ must produce a collective decision over a set of alternative social situations A $\{a_1, ..., a_m\}$. Each individual has a preference ordering P_i of the alternatives (this could be a simple ordinal ranking or a cardinal utility function). The group preference ordering, or social choice function, is some function

W, such that W represents the preferences of the group:

$$P_I = W(P_1, P_2, ..., P_n) \tag{1}$$

In Harsanyi's formulation of social choice, each individual's preferences are represented by a von Neumann-Morgenstern cardinal utility function $u_i(a_j)$ that obeys the standard axioms of Bayesian rationality [16]. Hence, equation (1) becomes:

$$U_I = W(u_1, u_2, ..., u_n) \tag{2}$$

In particular, W is defined to be the weighted summation of the individual utilities:

$$W(u_1, .., u_n) = \sum_{i \in I} \lambda_i u_i \tag{3}$$

where λ_i represents the weight given to agent i's utility function in the overall equation. To be socially rational, an individual maximises W over the different alternatives. This function represents how an individual agent may judge states of the world from a moral or social perspective. Harsanyi [8] has shown that his social welfare function satisfies the following important postulates. Firstly, the utility functions of all the individuals satisfy the axioms of rational behaviour under risk. Secondly, the social preference function also satisfies these axioms. Finally, if all individuals are indifferent between any two alternatives, from the viewpoint of their personal preferences, then an individual j will be indifferent between them from the standpoint of its social preference function. This provides an invaluable link between conceptions of individual and social rationality.

Our work adopts the basic mechanism of social choice defined in equation (3) (although see [12] for variants) as the means of ensuring that individual agents make socially responsible decisions. Using this formulation of social welfare, an individual has two separate preference orderings: its individual preferences (indicating an agent's own preference structure); and its social preferences (representing what an agent believes is best for the society in which it is situated). The task of analysing how the agent decides what action to perform can be formulated by differentiating how the agent makes decisions in terms of what it expects to get out of the action itself and what effect the action has on others. The effect that an action has on other agents can be determined by examining their utility functions. This is a feasible method for agent systems consisting of tens of agents (which is our primary target. However, when the number grows to hundreds or even thousands, agents will have large acquaintance sets and the calculation of all the necessary utility functions will be computationally expensive. In such circumstances, the agent will need to control this calculation by considering how resource bounded it currently is (see section 5 for a further discussion of this point). A further potential difficulty with the approach we have advocated stems from the inclusion of the acquaintance's utility functions. In an ideal world, each agent would know the utility function of all the other agents. However, in practice this is infeasible. Thus, in the formulation given here each agent uses an approximation of what it believes the other agents' utility functions to be. In terms of the balance between individual and social needs, equation

(3) can be re-written as:

$$W_i(a_j) = \lambda_i u_i(a_j) + \sum_{k \in \{I-i\}} \lambda_k u'_k(a_j) \tag{4}$$

By setting the weighting of a specific utility function to zero, an agent can eliminate the influence of that acquaintance on its decision making. For example, setting the agent's own weighting λ_i to zero removes any personal utility benefit consideration from its decisions. In this way, the agent can tailor its function to give more weighting to either individual or social concerns. Thus we can say at a coarse level, equation (4) becomes:

$$W_i(a_i) = \lambda_i u_i(a_j) + \lambda_{soc} \sum_{k \in \{I-i\}} u'_k(a_j) \tag{5}$$

where λ_{soc} is the weighting given to the social utility part of the function. To gain a perspective on the feasibility of using such a mechanism in practical multi-agent systems, the above decision making function has been implemented in a multi-agent testbed and evaluated as follows.

2.2 The Experimental Domain

The decision making framework was evaluated by applying it to the Phoenix fire fighting multi-agent simulation [5]. Here, the park is policed by several firebosses, each of which is responsible for a specific geographic region. Firebosses control a number of firefighting resources (bulldozers), that can be exchanged as and when necessary. The goal of the firebosses is to prevent fires claiming too much land. Hence, the success of an agent, and the utility it achieves, is measured as a function of land loss. The system's overall utility can also be measured in this way (i.e. land lost in the whole park). In order to use the framework described in section 2.1, we assume that the agents' utility functions are comparable[1]. A fireboss can estimate, given the current environmental conditions, how a fire is likely to spread and it can therefore choose the most effective allocation of its firefighting resources. These projections provide the agent with a means of determining the likely outcomes of different courses of action. Thus, for example, an agent can ascertain whether asking for a loan of resources from another fireboss is likely to be profitable. This choice is affected by the uncertainty of the reply of the other agents, and hence the decision maker will use a probability distribution based on whether the agent believes that i) there is a fire in the other agent's part of the park; and ii) whether the other agent is likely to give resources if it is possible for it to do so.

In what follows, we assume all agents are motivated towards the goal of reducing the amount of land lost as a result of fires. If an action results in large land loss, then that action has lower utility in comparison with an action that

[1] This is acceptable if we assume that the agents' utility functions are comparable up to an affine transformation [11].

produces lower land loss. Attaining more resources allows the agent to complete fighting the fire more quickly and so is preferred to simply fighting the fire with its current resources. However, asking for extra resources from another fireboss decreases that fireboss's effectiveness (and utility) when it is faced with a fire of its own. Therefore agents need to be able to estimate the likely utility of a particular resource distribution to their acquaintances in terms of predicted land lost in their region. Agents do this by assessing whether a fire is likely to occur in the acquaintance's area of the park and what effect particular resource distributions will have for the agent in this context (based on the acquaintances estimated utility function). This is then used in equation (5) to determine which action is socially preferred.

To illustrate the nature of social decision making in Phoenix, consider a scenario involving three agents (A, B and C)[2], each of which has control over two bulldozers. If a fire breaks out in A's area, A must decide whether to ask B or C for extra resources or proceed to fight the fire with its current resources. By obtaining an extra bulldozer, A will probably reduce the amount of land it loses and hence increase its utility. However, taking a bulldozer from B or C reduces their firefighting power and hence decreases their expected utility in the event of a fire. In addition, sharing bulldozers involves overheads based on the time it takes to communicate, as well as the delay of waiting for the extra resource(s) to arrive. Furthermore, A is not certain that its request will be granted, hence time may be wasted if its request is refused. Against this background, A's decision can be formulated in the following manner:

$$W(ask_C) = \lambda_A u_A(ask_C) + \lambda_{soc} \sum_{k \in \{B,C\}} u'_k(ask_C) \tag{6}$$

where $W(ask_C)$ represents the social welfare function of agent A asking agent C for a loan of resources, λ_A represents the weighting A gives to its own utility function and λ_{soc} is the weighting given to the social part of the utility equation. The utility function captures the domain specific considerations detailed above, such as the likelihood of success and the communication overheads involved in requesting a bulldozer. Each agent uses a discrete probability distribution to represent the likelihood that a resource request is accepted. Initially all agents use the same distribution, but these are updated over time as a result of interaction and resource exchange between the agents (e.g. when resource requests are accepted or rejected). Agent A will compute similar formulations for its other options (ask B and fight the fire with current resources) and then select the action that maximises the social welfare.

3 Evaluating the Effectiveness of Social Decision Making

Our aim is to evaluate the effectiveness of different decision making attitudes under varying levels of resource pressure. To this end, the experimental control variables are:

[2] For reasons of clarity, the example scenario considers just 3 of the systems' agents

– The utility function weightings λ_i and λ_{soc} (from equation (5)). These values can be altered to implement a wide range of decision making strategies. Here we use the following values:
  - ◇ 0 : 1 *Selfless*: care nothing about their own utility. These agents will not ask for help, but will grant any resource request that they receive (provided they are able to do so);
  - ◇ 1 : 0 *Selfish*: care only about their own utility. They will ask for resources from others, but will not loan out their own resources;
  - ◇ 0.5 : 0.5 *Balanced*: place equal concern on individual and social needs;
  - ◇ 0.25 : 0.75 *Social tendency*: are more social than selfish;
  - ◇ 0.75 : 0.25 *Selfish tendency*: are more selfish than social.

– The number of firefighting resources each agent controls (1, 2 or 3 bulldozers. In this context, 3 represents as many as the agent ever needs to fight a single fire);

– How fast each fireboss can make a decision (slow, medium or fast).

Resource pressure is exerted by: (i) altering the number of resources an agent controls and (ii) manipulating how quickly an agent can make its decisions. A given simulation run involves specifying when and where fires occur, how large they are initially, and what the environmental conditions are (wind speed etc.). Each run is randomly generated, involving between 2 and 6 fires. Statistically significant results are obtained by averaging over 12 runs. There are six firebosses in these experiments; although this number represents a fairly small multi-agent system, it nevertheless provides a reasonable indication of the feasibility of our decision making mechanism.

Two broad types of experiment were carried out. Firstly, ones where all agents used the same decision making attitude (section 3.1). This is to ascertain the performance profile of the various attitudes under varying resource constraints. Secondly, those in which agents have different attitudes (section 3.2). This is to ascertain the robustness of the decision making attitudes against a range of heterogeneous opponents. In all the experiments, individual utility is measured as a function of land lost and system utility is measured as a function of the land lost in the entire park. The latter is calculated as the aggregation of the land lost by all the firebosses.

3.1 Homogeneous Agent Societies

These experiments seek to test the following hypotheses about social decision making:

1. *The performance of the self-biased attitudes (selfish, selfish tendency) will deteriorate, both at the individual and system level, as the number of resources in the system are reduced;*
2. *The performance of the society-based attitudes (selfless, balanced and social tendency) will slowly degrade as resources are reduced, but not as much as the self-interested types;*

3. *Balanced agents will achieve the best system level performance since their decisions are based upon the welfare of the whole system.*

Figure 1 shows how the average individual performance of the different agent strategies is affected when resource pressure is exerted on the system. Runs using different rates of thinking (the other form of resource pressure) produce similar results, but because of space limitations are not shown here. It can be seen that self-biased agents, as predicted in hypotheses 1 above, perform poorly in times of scarce resources. This is because they waste time trying to obtain extra resources when i) it is perhaps unnecessary and ii) the system is comprised of agents of the same type that will not provide assistance. The selfish tendency strategy is better than the purely selfish one because requests for extra resources are only made if they are likely to have a noticeable benefit. The surprising result is that agents with a social tendency perform very badly in times of scarce resources. This is due to the fact that in times of scarce resources when resources are more valuable, these agents are likely to ask for help since the value of those extra resources outweighs any social considerations it has. However, other agents are also less likely to loan out their resources as it is too costly. This introduces a delay in the fire fighting and so means more land is lost. Balanced agents perform well since all utilities are considered equally and so the costs of asking for resources and loaning them out play an important role in the decision. This means balanced agents will ask for and loan out resources, but only if it is clearly beneficial. In times of plentiful resources, the performance of the different types becomes less disparate since agents generally have sufficient resources to minimise the impact of social interchange.

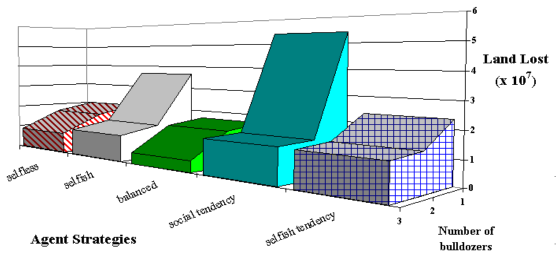

Fig. 1. Individual Agent Performance.

Figure 2 shows the cumulative land loss of the entire system. Here, agents with social tendencies generally perform well as they explicitly attempt to assess the system wide implication of their choices. However, selfless agents do not perform as well as the balanced or social tendency agents. This is due to the fact that selfless individuals miss the opportunity of attaining available resources from elsewhere (cf. balanced/social tendency attitudes) and hence poor individual performance accumulates into poor system performance. This cumulative effect also accounts for the extremely poor system level performance of the self-interested strategies. The selfless strategy does however represent an improvement on the self-interested strategies as it does not waste time asking for resources since taking a bulldozer away from another fireboss reduces their chances of fighting a fire effectively.

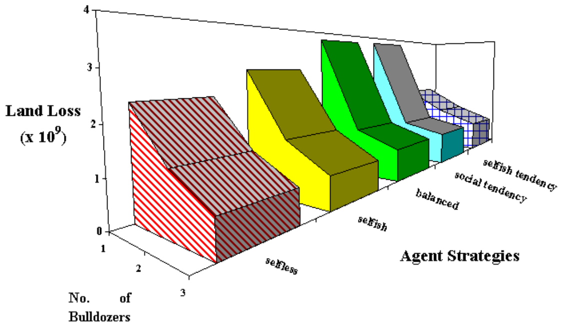

Fig. 2. System Performance.

3.2 Heterogeneous Agent Societies

To investigate the performance of a system comprising of agents using different strategies, the runs described for homogeneous societies were repeated using different percentage mixtures of the various strategies. In particular, different percentages of selfish agents (25%, 50% and 75%) were introduced into societies of the other decision making attitudes with the resource pressure kept at a constant level. We are especially interested in the impact of selfish agents since these should have the greatest detrimental affect on the performance of socially rational societies. To this end, we wish to explore the following hypotheses:

4. *The performance of selfless agents will decrease rapidly as more selfish agents are introduced;*
5. *The performance of balanced agents will be resilient to the introduction of selfish agents;*
6. *Mixing selfish and socially motivated agents in one society will sometimes produce system performance that is superior to that of the homogeneous societies of either type.*

Figure 3 shows two graphs, *(a)* showing both the average individual performance of selfless and selfish agents in societies in which the percentage of selfish agents is steadily increased and the other *(b)* a similar graph, but with societies that are a mixture of balanced and selfish agents. The individual performance of both types suffer as the number of selfish agents is increased. However, the balanced agents are less susceptible to the increase in the number of selfish agents since they have an inbuilt concern for their own utility as predicted in hypothesis 5 above. This means they will not unquestioningly give resources to others if they can profit from retaining them.

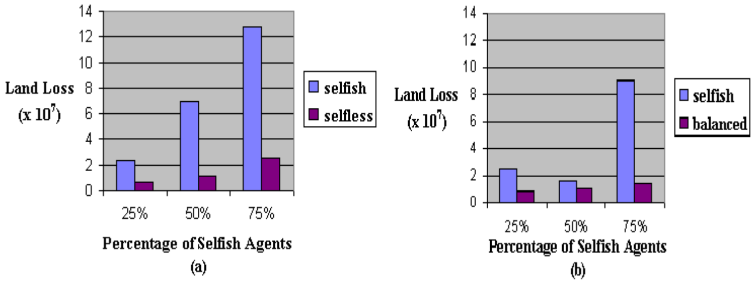

Fig. 3. Individual Performance in (a) Selfish/Selfless and (b) Selfish/Balanced Societies

Figure 4 demonstrates how the mixed societies perform on a system level. The gradual introduction of more selfish agents decreases overall system performance for both mixtures. However, the society consisting of balanced agents shows a more steady decline in performance than the society containing selfless agents. Again this occurs because balanced agents are more concerned for the overall system and not just for individual or altruistic concerns. One point to note is the initial performance improvement of the selfless/selfish society. When there is a small number of selfish agents, and several selfless agents willing to accede to requests, overall performance improves since resources in the system are being distributed more effectively than would be the case if the system consisted solely of selfless agents. This can be related to hypothesis 6 where we expected that system performance would actually improve with some mixtures of agents. As the number of selfish agents increases, there are fewer opportunities for these agents to gain resources, so performance again deteriorates.

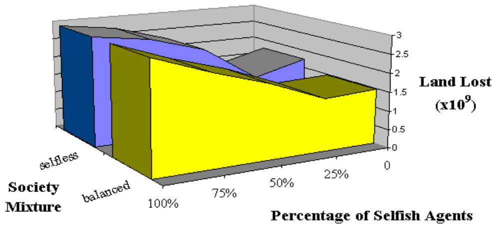

Fig. 4. System Performance of Heterogeneous Societies.

4 Related Work

Agent rationality has been widely debated and studied in the field of agent research [15]. Decision theory [18] has emerged as the definitive descriptive and normative theory of rational decision making. The fundamental principle of decision theory is the maximisation of the agent's utility function under certain axioms of uncertainty and utility [13]. Game theory is also concerned with the rational behaviour between two or more interacting individuals [17]. Each agent has a payoff or utility function that they attempt to maximize based on the information they have about the strategy of the other individual(s). This payoff function represents the preferences of the individual, though it can be based on altruistic motives in the case where more global/social concerns are the dominant philosophy. There are, however, a number of problems with game theory with regards to the social aspects of decision making. One is the inability to deal adequately with some social notions such as cooperation. In fact, without the introduction of some binding force ensuring cooperation, the theory can produce sub-optimal results, as shown by the prisoners dilemma example. Furthermore, although both game and decision theory provide simple and attractive formalisms of individual action choice, they have been criticised on the grounds that they reveal nothing about the motivations of the agents making the decisions [2]. For example, both disciplines can produce socially acceptable results if the utility functions used incorporate some social information, but these theories provide no answers as to *how* this can be done or even *why* this should be done. This, in turn, is of little use when attempting to understand, describe and ultimately build socially aware agents. Thus, we adopt some fundamental principles of these theories but expand these ideas to explore our ideas of social reasoning.

A consistent theme in the work of Castelfranchi [1, 3] is the concept that sociality is derived from the individual mind and social action. Social rationality, and in particular an agent's social power, is described via manipulation of dependence relationships between agents. Agents may interfere, influence and adopt goals of their acquaintances as a result of the manipulation of these relationships. Such notions can then form the basis of a variety of social actions. Although underlining the need to explore and emphasise the social aspects of an agent's make-up, this line of work addresses the philosophical rather than practical questions of how this should be achieved. Building on this, Cesta *et al.* [4] explore the practicalities of social decision making by experimenting with a variety of social attitudes. Their work mainly covers simple, rather rigid, agent systems and concentrates on how the introduction of exploiters into a society effects system performance. Their results tie in with our findings regarding the introduction of what we have called selfish agents. When exploiters (selfish agents) are introduced into their system, the performance of the system decreases, a result which is magnified as resource pressure is exerted. Although they look at some effects of resource bounds, this is not the main thrust of the work. Also, there is no discussion of how individual autonomy is balanced with social concerns in such contexts.

Jennings and Campos [9] define a social equivalent of Newell's conceptualisation of individual agent rationality that they term the Principle of Social Rationality. Social rationality is defined as the action choice of an individual based on global concerns. To add substance to this definition, Kalenka and Jennings [10] describe several social attitudes that can be ascribed to agents under this principle. Their work provides a framework for defining the different social attitudes that an agent may possess, including helpfulness and cooperativity. However, the missing element in their work is the practical consideration of resource bounds on the performance of social agents. Their framework also restricts the level of analysis that can be performed with regards to an agent's different relationships in the society. For instance, there is no mechanism to employ when the agent finds itself as a member of multiple groups or coalitions.

More socially minded decision making attitudes have been investigated in the socio-economic literature under the umbrella of social welfare functions (also collective choice rules or preference aggregation rules) [11]. Here the main emphasis is on how a group of agents can collectively make decisions. The decision maker can either be several agents making a joint decision or an individual making a decision that has global consequences. These functions have been shown to have the advantage of Pareto optimality, but have the disadvantage that equity is not preserved in the group (i.e. the decision is not fair to everyone, for example in the case of the distribution of wealth). There are also concerns as to how the utility functions are derived and how they should be combined in an overall function to reflect group choice. These issues are also important when we consider software agents and, at present, there are no comprehensive solutions to these problems. However, we do believe that practical assumptions can be made about the origin and structure of the utility functions used by agents, and that with further experimentation into these issues, useful insights can be found.

5 Conclusions and Future Work

This paper has outlined the case for a more socially aware approach to decision making in a multi-agent context. A novel agent decision making framework, incorporating insights from work on social welfare functions, has been devised to tackle the problem of social decision making by an individual in a multi-agent environment. This framework provides a means of describing and analysing how an individual agent may approach the task of making socially acceptable decisions in a multi-agent system. More importantly, perhaps, is the empirical demonstration of the effectiveness of various socially aware decision functions in a range of problem solving scenarios. Our results indicate that decision attitudes based on social concerns perform better in resource bounded contexts than the more traditional, self-interested attitudes. In particular, our results for balanced agents demonstrate the importance of considering both the individual and system consequences of decision making. Furthermore, this work investigated the effect of having several different decision making attitudes in the same system. Here again we highlighted the importance and effectiveness of basing decisions on both individual and social concerns by demonstrating the robustness of balanced agents in the face of exploitation by selfish agents. These experiments also demonstrate the importance to the performance of the individual and the system, of the mixture of strategies used by the participating agents and how this must be considered in the individual's decision making function.

The next step is to utilise our empirical findings to construct resource bounded, socially rational decision makers. Here, the empirical data provides the basic knowledge about social decision making. The results indicate when the various attitudes are effective (both from an individual and system perspective). This know-how can now be used as a basis for the meta-reasoning component of a resource bounded reasoner. In particular, we are currently embedding the key findings into the control level of our Phoenix agents. We envisage that our agents will use the basic social decision function described in section 2, but to the address problems of resource bounded behaviour they will: i) dynamically modify the weightings given to the respective parts of the social welfare function; and ii) alter the amount of computation performed by modifying the acquaintances used in the calculation. This will provide adaptive behaviour over different resource contexts. Thus, for example, the agent may weight its own utility higher than others and only include a subset of its peers in the social utility calculation in times of heavy resource pressure and when it believes it can obtain resources from others. However, when the agent has more time or resources to play with, then it may fully consider all parts of the social welfare function when deciding what to do.

In terms of extending our work, we need to further investigate how agents can dynamically build relationships with one another and then use this knowledge to learn how to operate more efficiently. For instance, an agent may learn to ask certain acquaintances for help and not others (e.g. if it has received negative responses to previous resource requests). Such interaction adaption can also help agents choose which acquaintances they should consider in the social part of their utility equation. This is especially useful when the agent is faced with heavy time pressures since it need only perform calculations for the acquaintances that it deems important. This adaption can also be used to consider what effect the appearance of coalitions or sub-groups has on the agent's decision making strategy.

Acknowledgements

The authors would like to thank all the reviewers of this paper for their helpful comments, especially Cristiano Castelfranchi for his constructive feedback and thought provoking discussions.

References

1. C. Castelfranchi. Social power a point missed in multi-agent, DAI and HCI. In Y. Demazeau and J.P. Muller, editors, *Decentralised AI.*, pages 49–62. Elsevier Science Publishers, 1990.
2. C. Castelfranchi. Limits of economic and strategic rationality for agents and MA systems. *Robotics and Autonomous Systems*, 24:127–139, 1998.
3. C. Castelfranchi. Modelling social action for AI agents. *Artificial Intelligence*, 103:157–182, 1998.

4. A. Cesta, M. Micheli, and P. Rizzo. Effects of different interaction attitudes on multi-agent system performance. In W. Van de Welde and J. Peram, editors, *Proc. of MAAMAW 96*, pages 128–138. Springer Verlag, 1996.
5. P.R. Cohen, D.M. Hartland M.L. Greenberg, and A.E. Howe. Trial by fire: Understanding the design requirements for agents in complex environments. *A.I. Magazine*, 10(3):32–48, 1989.
6. J. Doyle. Rationality and its role in reasoning. *Computational Intelligence*, 8(3):376–409, 1992.
7. B. Grosz. Collaborative systems. *A.I. Magazine*, 17:67–85, 1996.
8. J. C. Harsanyi. *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations.* Cambridge University Press, 1977.
9. N. R. Jennings and J. Campos. Towards a social level characterization of socially responsible agents. *IEE Proceedings on Software Engineering*, 144(1):11–25, 1997.
10. S. Kalenka and N.R. Jennings. Socially responsible decision making by autonomous agents. In *Proceedings of the Fifth International Colloqium on Cognitive Science (to appear)*, 1999.
11. K.P.Corfman and S. Gupta. Mathematical models of group choice and negotiations. In J. Eliahberg and G.L. Lilien, editors, *Handbook in OR and MS*, pages 83–142. Elsevier Science Publishers, 1993.
12. J. F. Nash. The bargaining problem. *Econometrica*, 28:155–162, 1950.
13. S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach.* Prentice Hall, 1995.
14. S. Russell and E. Wefald. *Do The Right Thing: Studies in Rationality.* MIT Press, 1991.
15. Stuart Russell. Rationality and intelligence. *Artificial Intelligence*, 94(1):55–57, 1997.
16. L. J. Savage. *The Foundations of Statistics.* Wiley, New York, 1954.
17. T. C. Schelling. *The Strategy of Conflict.* Harvard University Press, Cambridge Massachuttes, 1963.
18. J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behaviour.* Princetown University Press, 1944.