Simple Negotiating Agents in Complex Games: Emergent Equilibria and Dominance of Strategies

Peyman Faratin[†] Mark Klein[†] Hiroki Sayama[‡] Yaneer Bar-Yam [‡]

[†] Center for Coordination Science, MIT, Cambridge, MA, 02139, USA
{peyman,m_klein}@mit.edu
[‡] New England Complex Systems Institute, 24 Mt. Auburn St., Cambridge, MA 02138, USA
{sayama,yaneer}@necsi.org

Abstract. We present a simple model of distributed multi-agent multi-issued contract negotiation for open systems where interactions are competitive and information is private and not shared. We then investigate via simulations two different approximate optimization strategies and quantify the contribution and costs of each towards the quality of the solutions reached. To evaluate the role of knowledge the obtained results are compared to more cooperative strategies where agents share more information. Interesting social dilemmas emerge that suggest the design of incentive mechanisms.

1 Introduction

Arguably one of the most significant contributions of AI to the theory of computation has been models of the relationship between knowledge, computation and the quality of solution (henceforth referred to as the K-C-Q relationship) [24, 18]. In this article computation is equivalent to the process of searching a space of possibilities [21]. AI has shown that there exists a hierarchy of tradeoffs between K, C and Q, with models that achieve perfect optimal results but at the cost of requiring omniscience and unbounded agents, to models that sacrifice optimality of Q for a more realistic set of requirements over K and C [22]. Different agent architectures are then entailed from different K-C-Q relationship theories. However, the rise of Distributed Artificial Intelligence (DAI), and Multi-Agent Systems (MAS) in particular, has given rise to the need for new models of this relationship, of not only the *local* K-C-Q relationship for the agent itself but also the *social* K-C-Q relationship. In social systems, the underlying relationship is regulated through the protocol of interaction.

Game theoretic models have emerged as the most popular candidate social models in MAS [15, 20, 23]. However, in such prescriptive models desirable solutions are achieved at the cost of make very restrictive assumptions over both K and C [4, 10, 6]. In the types of problems that interests us K and C are limited because not only do agents have an incomplete knowledge of their own utility functions but they also have, at best, imperfect and at worst no knowledge of the others' utility functions. We say that an agent's knowledge of its own utility function is incomplete when it can not compare the utility of a point in the utility landscape with all the other points and searching this landscape exhaustively is computationally prohibitive. Additionally, agents rarely have full knowledge of others' utility landscape. The best an agent can do is to reason with imperfect knowledge by forming approximations of others' utilities (and later correcting for imperfections using some update rule such as Bayesian rule [19]). In the worst case, agents have no knowledge of the others utility landscape. In either case, lack of perfect knowledge results in a process of negotiation that requires agents to search for mutually acceptable solutions. Therefore the problem is how to design search algorithms that are not computationally prohibitive but also result in good solutions given limited knowledge.

Descriptive (or *evolutive*) models of evolutionary game theory are emerging as alternatives that make fewer restrictive assumptions as well as providing models of the dynamics of interactions [25, 1, 4]. In these models the individual *is* merely a strategy which is subjected to survival criteria in a population of other strategies. The problems associated with the prescriptive models are eliminated by replacing the agents with simple stimulus-response machines. Under this methodology coordination *emerges* from interactions between simple agents. In this paper we adopt a similar methodology as the evolutive approach to address the computation and knowledge complexities involved in the problem of large scale and interdependent multi-issued contract negotiation between autonomous agents in an open system. In particular, we show that distributed decision making using two simple approximation methods (the hillclimbing and the simulated annealing) can result in interesting social and local dynamics when the knowledge, computation and the quality of the solution is regulated by a voting protocol.

Section 2 is a brief introduction and overview of the contract negotiation problem and its features. We then describe a mediated negotiation protocol in section 3 followed by a model of the problem objective function in section 4. Sections 5 and 6 then presents a pair of simple selfish (local and hence knowledge poor) and cooperative (global and hence knowledge richer) evaluation strategies respectively. In the penultimate section 7 we presents the simulation results. Finally, section 8 identifies the weaknesses and the future directions of research.

2 The Contracting Problem

The particular complexity of interest for us is the knowledge and computation involved in open system multi-party contract negotiation over multiple clauses (we refer to a contract clause as an issue). In particular, we are interested in design of distributed agent decision algorithms that select a negotiated outcome that are both individually and socially good given that the agents are bounded in both knowledge and computation. In multi-issue negotiation computation is often bounded because of *scale* and *issue dependencies*. Scale is a problem because even a relatively simple contract can easily involve 100 issues. Even with only two alternatives for each issue, the size of the search space is roughly 10^{30} possible contracts, too large to be explored exhaustively. This implies that agents will not know which contracts are best for them a priori. Therefore they have to search the contract design space. The problem of issue dependency occurs for the following reason. If the issues being negotiated over are assumed to be independent (i.e. the utility to an agent of a given issue choice is independent of what selections are made for other issues), then the utility function that aggregates the individual utilities is a linear one (often additive). This assumption significantly simplifies the agents' local decision problem of what issue values to propose in order to increase their own utility: they simply need to ”hillclimb” in a straight line up the utility slope. Almost all work to date on computational negotiation algorithms has in fact focused on the independent issue case [11, 9]. Real world contracts, however, are highly inter-dependent. For example, in a supply chain the output of one sub-step in a process will often have to match the input of the next sub-step, so the utility of one choice is highly dependent on the other choice. When issue interdependencies exist, the utility function for the agents exhibits multiple local optima, and the process of finding a satisfactory contract becomes much more complicated because it will often make sense during the search process to consider contracts with low utility while moving towards contracts with potentially much higher utilities.

Additionally, in open systems knowledge is scarce because of information exchange constraints. MAS negotiation approaches do not assume that agents will be cooperative and make choices based on their impact on global utility, but rather that they will be purely self-interested [8]. This makes negotiation well-suited to open systems where the participants can be diverse and we can not assume that all of them will be cooperative or even benevolent [10]. Because of the competitive nature of negotiation, however, agents will typically wish to reveal as little information as possible about themselves to other agents. Therefore, computation is knowledge poor.

Another consequence of the open system assumption underlying negotiation approaches is that, since we can not assume that agents will be altruistic, we must design negotiation protocols such that the individually most beneficial negotiation strategies also produce the greatest social welfare (i.e. the greatest aggregate utility summed over all agents) [20, 23]. In other words, we want the socially most beneficial strategy to also be the individually dominant one so that most agents will tend to use it. There has been significant success achieving these property in the independent issue case, but the specification of dominant strategies becomes more complicated, as we shall see, for the interdependent issue case.

The challenge therefore is to define contract negotiation algorithms that are suited to large design spaces with interdependent issues but are sensitive to the information exchange constraints and strategy dominance concerns important in open system contexts. Approximation techniques to such hard problems have been successful in the past for other such hard problems such as machine vision [3]. We present agent strategies as an approximate optimization algorithms in sections 5 and 6 after an informal description of the protocol of negotiation below.

3 The Negotiation Protocol

A contract S, is an N dimensional boolean vector where $S_i \in \{-1, +1\}$, represents the presence or absence of a “contract clause” i. In the simulations the size of the vector S was set to 100 bits, corresponding to a space of 2^{100}, or roughly 10^{30} possible contracts.

The contract search policy is encoded in the negotiation protocol. Because generating contract proposals locally is both knowledge and computationally expensive (see [11] for an approach to this problem) we adopt an indirect interaction protocol between two agents by delegating the contract generation process to a centralized mediator [12]. The mediator proposes a contract S^t at time t. Each agent then votes to accept or reject S^t. If both vote to accept, the mediator iteratively mutates the contract S^t and generates S^{t+1}. If one or both agents vote to reject, a mutation of the most recent mutually acceptable contract is proposed instead. The process is continued until the utility values for both agents become stable (i.e. until none of the newly contract proposals offer any improvement in utility values for either agent). In this manner, search of the joint and local utility landscapes of the agents is iteratively carried out by the mediator who incrementally modifies *a* randomly selected "contract clause" from a *present* state $(+1)$ to a *absent* state (-1) or vice versa. Agents then decide whether to accept or reject this new single "step" in their local utility landscape. Note that this approach can straightforwardly be extended to a N party (i.e. multi-lateral) negotiation.

4 The Objective Functions

At the end of each modification agents must vote to either accept or reject a contract. This decision is based on the utility of a contract. The utility is defined as the linear combination of all the pairwise influences between issues. That is:

$$u = \sum_{i,j} J_{ij} S_i S_j \qquad (J_{ij} = J_{ji}, J_{ii} = 0) \tag{1}$$

where u is the local contract utility of an agent, and $J_{ij} \in [-1, +1]$ is the utility influence matrix between issues i and j, which represents the utility increment or decrement caused by the presence of a given pair of issues. We assume that this pairwise effect is symmetric and not defined for $i = j$, as shown in parentheses. Each negotiation agent has a local and private J_{ij} matrix which is different to that of all other agents. We do not include the increment or decrement by the presence of a particular clause itself, since they are linear terms that do not affect the non-linearity of the utility space that is of our main interest. Also note that interactions among three or more clauses are omitted for simplicity, since the search is already complex enough for problems involving hundreds of pairwise relations. The key issue we want to model is the conflicts between agents, and they are guaranteed in this simple model when the objective function is defined with different distributions of J_{ij} for different agents.

Given the utility of a contract agents must then decide in a distributed manner whether to accept or reject a contract S^t at time t. To examine the role of knowledge (or willingness to share information) on the quality of solution in this model of negotiation we distinguish two types of evaluation strategies, local and global. These strategies, described in more detail in sections 5 and 6 respectively, have counterparts in game theory [13].

5 Local Agent Strategies

In a local evaluation strategy agents do not share any information and evaluate the generated contract S in a distributed or local fashion. Thus, what is being optimized is the *individual* utility of the agent u given in equation 1. Specifically, agents' decision to accept or reject a contract S^t is a function of the magnitude of the difference between S^t and S^{t-1} defined as $\Delta u^k(S^t, S^{t-1}) = u^k(S^t) - u^k(S^{t-1})$, for agent k at time t.

Two computationally inexpensive decision algorithms were evaluated: a hillclimber and a simulated annealer. A hillclimber decision rule is defined as:

$$p^k(S^t) = \begin{cases} 1 & \text{if } \Delta u^k(S^t, S^{t-1}) > 0.0 \\ 0 & otherwise \end{cases} \tag{2}$$

where $p^k(S^t)$ is the probability that the agent k will accept the given contract S^t generated from the previous contract S^{t-1}. Therefore, a hillclimbing search evaluation function only accepts a contract if and only if the utility of the contract S increases monotonically when *an* issue is changed.

However, the formulation of the objective function given by equation 1 has multiple optima. Therefore, this steepest ascend algorithm is known to be incapable of escaping local maxima of the utility function. Search success can be improved by adding thermal noise to this decision rule [14]. The concrete probabilistic rule used to simulate thermal noise is:

$$p^k(S^t) = \begin{cases} 1 & \text{if } \Delta u^k(S^t, S^{t-1}) > 0.0 \\ e^{\Delta u^k(S^t, S^{t-1})/T} & otherwise \end{cases} \tag{3}$$

where $p^k(S^t)$ is the probability that the agent k will accept the given contract S^t at temperature T. The policy of decreasing T with time is called simulated annealing [14]. The success of the rule is based on its elimination of barriers between local minima by giving a small chance to large utility configurations. This decision rule is known to reach utility equilibrium states when each issue is changed with a finite probability and time delays are negligible [14].

Like evolutive models these decision rules completely specify the agents. We categorize these rules in to selfish strategies (equation 2) accepting S^t iff $u(S^t) > u(S^{t-1})$, or cooperative (equation 3) by accepting a contract S^t when $u(S^t) \leq u(S^{t-1})$ with a certain probability determined by the thermal noise. Note, that if we define accept and reject as the only *pure* strategies, then both the hillclimbing and simulated annealing profiles are equivalent to implementing a *mixed* strategy of accept *and* reject, one deterministically (hillclimber) and the other according to a probability determined by the thermal noise [13]. Furthermore, although not explicitly modeled the thermal noise variable can be interpreted as encoding in to the behaviour of agent its willingness to take risks in return for future possible gains. The higher T the more risk loving and, conversely, the lower T the more risk averse the agent.

6 Coalition Strategies

In the above section agents' evaluation of S^t is distributed and local. Global evaluation strategies were also considered to assess the role of knowledge on the quality of computation and compare the results obtained from distributed v.s centralized decision making (the latter approach is the most popular with OR algorithms). We call this global evaluation a coalition strategy because agents have *agreed* prior to the game to a strategy of sharing their local utility for a contract and optimizing some social welfare function (simply defined as the sum of the local utilities) during the game.[1] This formulation is equivalent to agreements made before a game on a pair of strategies that maximizes the joint social welfare for two player games in cooperative game theory [16]. Local strategies are equivalent to non-cooperative games where there is no pre-negotiation agreements before the ensuing game [17]. Note that whereas there are a total of 2^N possible pairs of combination of hillclimbing simulated annealing strategies for N agents, in cooperative evaluations there are only two strategies available, hillclimbing or simulated annealing. As will be shown quantitatively later a coalition removes the gaming part of the interaction.

7 Simulations: Procedures and Results

Simulations were run with two agents referred to as a and b. The contract length N was set to 100 where each bit was initialized to a value $\{-1, +1\}$ randomly with a uniform distribution. The utility influence matrices J_{ij}^k for agent k was also randomly selected to be in the interval $[-1, +1]$ with a uniform distribution. Different influences matrices were used for each simulation run, in order to ensure our results were not idiosyncratic to a particular run. The initial temperature was set to 10 and decreased in steps of 0.1 to 0. Final average utilities were collected for 100 runs for each temperature decrement. Since there were only two agents the simulation pairing of local strategies making up the encounters in negotiation were: 1) a pair of local hillclimbers, 2) a pair of local simulated annealers and 3) one local simulated annealer and one local hillclimber (the assymetric case).

Figure 1 shows the results of the social welfare for different pairings of both local and coalition strategies. Relatively best results were observed when agents were cooperative and agreed on a joint annealing strategy to optimize the sum of their local utilities (filled circle) rather than optimize only their local utilities (all unfilled data points). However the final benefits of a coalition was lower when the optimization strategy was hillclimbing (filled triangles up) than a pair of local annealing strategies (unfilled circles), even though initially hillclimbing cooperatively over the joint utilities leads to better results quicker. These observations, and in general comparisons of the annealers and hillclimbers independently of the level of cooperation, suggests the natural C-Q tradeoff between search time and solution quality, where initially searching more (annealers at high temperatures), although costly, will in the long term result in better solution quality. This is confirmed by the observation that the presence of a local annealer agents always increases the final social welfare. The social welfare for the two local annealer case (unfilled circle in figure 1) was roughly 40% greater than that of the two local hill-climber case (unfilled triangle up in figure 1), and the assymetric case (unfilled diamond in figure 1) produced a smaller but still significant 15% improvement

[1] Note, this social welfare formulation is the pareto-optimal concept of efficiency [7]. However, unlike Nash bargaining solution [16] pareto-optimality, although efficient, does not imply fairness. Although we have some results on this issue we do not address the problem of fairness in the coalition in this paper.

over the local hill-climbers. The results also show that both local and coalition hill-climbers reach equilibrium sooner than the annealers. Local hill climbers for example typically reached stability after roughly 100 proposal exchanges, while the annealers approached equilibrium values after roughly 800 proposal exchanges. This is to be expected since the combination of hillclimbing search to either the global or local welfare is thresholded because hill climbers simply climb to the top of the *closest* utility optimum from the starting position and then stop, while annealers can, when at a high temperature at least, explore different multiple optima in the utility function. One might argue that the gain in utility achieved by using annealers was relatively unimportant given the time taken to achieve these results. However, the complexity of computation used to produce the observed quality of results is insignificant given the simplicity of the simulated annealers. This is demonstrated by noting the run times were on the order of seconds. Furthermore, although coalition hillclimbers were dominated by local

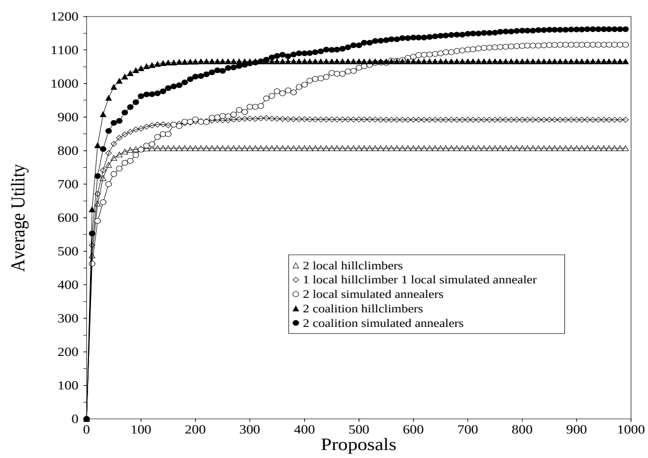

Fig. 1. Social Payoffs for both local and coalition strategies

annealers on the long term, the dominance of both coalition annealing and hillclimbing strategies over local strategies, shows that sharing information is generally better than not, thereby quantitatively confirming the K-Q relationship that more knowledge results in better solution quality. This relationship can be seen by comparing the social welfare obtained when agents make decisions in a coalition (filled data points) with decisions made locally (unfilled data points).

Individual payoffs were then examined to investigate the relationship of C-Q with local utility metric of Q. Figure 2 shows the observed individual payoffs for different pairs of local strategies. One observation is that if the other agent is a local hill-climber, an agent is then individually better off being a local hill-climber, but fares very badly as local annealer. If the other agent is an annealer, the agent fares well as an annealer but does even better as a hillclimber. The highest social welfare, however, is achieved when both agents are annealers. This pattern can be readily understood as follows. At

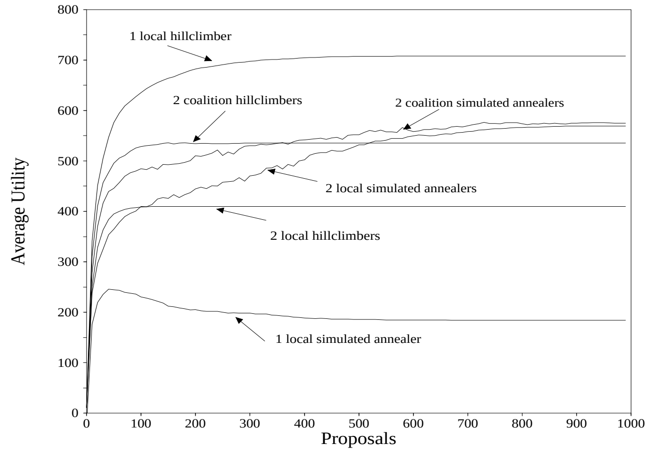

Fig. 2. Individual Payoffs for local and coalition strategies

high virtual temperature an annealer in the assymetric case accepts almost all proposed contracts independently of the cost-benefit margins. The asymmetry means that the co-operation of annealers *permits* more exploration, and hence arrival to higher optima, of hillclimber's utility landscape. However, this cooperation is not reciprocated by hill-climbers who act selfishly. Therefore, gains of hillclimbers are achieved at the cost of the annealer. Finally, the local benefits of coalition strategies closely reflect the same rank orderings as the social benefit results shown in figure 1. Figure 3 represents the

	Agent b Annealer	*Agent b Hillclimber*
Agent a Annealer	550/550	180/700
Agent a Hillclimber	700/180	400/400

Fig. 3. Payoff Matrix of the Game

underlying game as a matrix of final observed utilities for all the pairings of hillclimber and annealer strategies. The results confirm that this game is an instance of the prisoner's dilemma game [2], where for each agent the dominant strategy is hillclimbing. Therefore, the unique dominating strategy is for both agents to hillclimb. However, this unique dominating strategy is pareto-optimally dominated when both are annealers. In other words, the single Nash equilibria of this game (two hillclimbers) is the only solution not in the Pareto set.[2]

One possible solution to this problem is to anneal or hillclimb not over the individual utilities (the distributed case) but instead over the joint utilities (the coalition case,

[2] Note, for von Neumann and Morgernstern utility functions, the same underlying game structure is achieved through any affine transformations of the utility values.

figure 2). Then an institution or some centralized trusted third party eliminates gaming between the agents altogether by centrally executing a commonly known strategy, and searching through the set of possible contracts *given* agents local utility functions. However, truthful revelation of utility functions is a strong assumption and problematic and requires the design of additional mechanisms that implement the Revelation Principle where agents are incented to truthfully reveal their true utility functions [5]. Alternatively, agents may be permitted to evolve towards the socially dominant strategies of the underlying game incrementally [25]. In the context here this means that the design process of negotiating machines will eventually lead to design of machines that implement the simulated annealing strategy.

8 Conclusions and Future Work

The main contribution of this work has been to demonstrate how simple descriptive models of interactions can account for the dynamics of negotiation and lead to the emergence of interesting social outcomes. Two simple and strategic agent evaluation algorithms (hillclimbing and simulated annealing) were presented, regulated by a mediated voting protocol, for solving complex and large scaled problem of multi and interdependent issued contract negotiation. We have shown that there exists a tradeoff between the quality of solution reached and the computations involved with hillclimbing and simulated annealing algorithms. More search led to better contracts when annealing over either the individual or the joint utility landscape. Conversely, less search leads to relatively poorer contracts when hillclimbing over either the individual or the joint utility landscape but at a relatively lower costs. We have also shown the benefits gained on the quality of solution in cooperative negotiations where more knowledge is shared. Furthermore, we have demonstrated how local decision making can lead to the prisoner's dilemma games, a problem that is relevant to any distributed synthesis task involving negotiation (not just contract formation).

There are a number of future directions. Work is in progress to benchmark the computational complexity of the developed algorithms against centralised algorithms from Operation Research and evaluating their optimality against solution concepts such as the Nash bargaining Solution and Pareto-Optimality. Additionally, work is under way to develop a better causal model of the parameters (such as the starting temperature, the number of issues, the cooling regime, etc.) that regulate the K-C-Q relationship which can then support decision making over parameter values. Other, more intelligent, combinatorial optimization techniques are also being evaluated that better implement a contract generation policy. Finally, in the longer term we aim to develop and comparatively analyze additional agent strategies such as reciprocation, learning and evolutionary rules that execute in distributed or centralized repeated interaction protocols.

References

1. R. Axelrod. *The Evolution of Cooperation*. Basic Books, Inc., Publishers, New York, USA., 1984.

2. R. Axelrod. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton Press, Princeton, New Jersey., 1997.
3. D.H Ballard and C.M. Brown. *Computer Vision*. Prentice Hall, Englewood Cliffs, NJ, 1982.
4. K. Binmore. *Essays on the foundations of game theory*. Basil Blackwell., Oxford, UK, 1990.
5. K. Binmore. *Fun and Games: A Text on Game Theory*. D.C. Heath and Company., Lexington, Massachusetts, 1992.
6. C. Castlefranchi and R. Conte. Limits of strategic rationality for agents and MA Systems. In M. Boman and W. Van de Velde, editors, *Multi-Agent Rationality: Proceedings of the 8th European Workshop on Modeling Autonomous Agents in Multi-Agent World, MAAMAW'97*, number 1237 in Lecture Notes in Artificial Intelligence. Springer-Verlag, 1997.
7. G. Debreu. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*. Wiley, New York, 1959.
8. Edmund H. Durfee and Jeffrey S. Rosenschein. Distributive problem solving and multi–agent systems: Comparisons and examples. In *Proceedings of the 13th International Workshop on Distributed Artificial Intelligence,Seatle, WA*, pages 94–104, Seatle, WA, July 1994.
9. H. Ehtamo, E. Ketteunen, and R. Hamalainen. Searching for joint gains in multi-party negotiations. *European Journal of Operational Research*, 1(30):54–69, 2001.
10. P. Faratin. *Auotmated Service Negotiation between Autonomous Computational Agents*. PhD thesis, Department of Electronic Engineering, Queen Mary and Westfield College, University of London, 2000.
11. P. Faratin, C. Sierra, and N. R. Jennings. Using similarity criteria to make negotiation trade-offs. In *Proceedings of the International Conference on Multiagent Systems (ICMAS-2000), Boston, MA.*, pages 119–126, 2000.
12. K. Fisher and W. Ury. *Getting to Yes: Negotiating Agreement without giving in*. Houghton Mifflin, Boston, MA, 1981.
13. R. Gibbons. *A Primer in Game Theory*. Harvester Wheatsheaf, New York, 1992.
14. S. Kirkpatrick, C.D. Gelatt, and M.P. Vecci. Optimization by simulated annealing. *Science*, pages 671–680, 1983.
15. S. Kraus and D. Lehmann. Designing and building negotiation automated agent. *Computational Intelligence 11*, 11(1):132–171, 1995.
16. J. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950.
17. J. Nash. Non-cooperative games. *Annals of Mathematics*, 54:286–95, 1951.
18. J. Pearl. *Heuristics*. Addison-Wesley, Reading, MA, 1984.
19. H. Raiffa. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Addison-Wesley, Reading, MA, 1968.
20. J. S. Rosenschein and G. Zlotkin. *Rules of Encounter*. The MIT Press, Cambridge, USA, 1994.
21. S. Russell and P. Norvig. *Artificial Intelligence: A modern approach*. Prentice Hall, Upper Saddle River, New Jersey, 1995.
22. S. Russell and E. Wefald. *Do the Right Thing*. The MIT Press, 1991.
23. T.W. Sandholm. Distributed rational decision making. In G. Weiss, editor, *Multiagent Systems*, pages 201–259. The MIT Press, Cambridge, Massachusetts, 1999.
24. H. A. Simon. *The Sciences of the Artificial*. MIT Press, Cambridge, Massachusetts., 1996.
25. J. Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, 1982.