Using an Annealing Mediator to Solve the Prisoners' Dilemma in the Negotiation of Complex Contracts

Mark Klein[1], Peyman Faratin[1], and Yaneer Bar-Yam[2]

[1] Massachusetts Institute of Technology, NE20-336 Cambridge MA 02139
617 253-6796 (tel), 617 352-4424 (fax) {m_klein, peyman}@mit.edu
[2] New England Complex Systems Institute, 24 Mt. Auburn Street, Cambridge MA 02138
yaneer@necsi.org

Abstract. Research on computational models of negotiation has focused almost exclusively on defining simple contracts consisting of one or a few independent issues, implying tractable contract spaces as well as single-optima utility functions for the agents involved. Many real-world contracts, by contrast, are much more complex, consisting of multiple inter-dependent issues, resulting in intractably large contract spaces and multiple-optima utility functions. Complex contracts require negotiation algorithms substantially different than those that have been considered to date for simple contracts. Previous work by the authors has shown that endowing the negotiating agents with a time-decreasing willingness to provisionally accept inferior contracts (i.e. to 'anneal') results in superior final contract utilities, but creates a prisoners' dilemma wherein agents are individually incented not to anneal. In this paper we report a solution to this problem wherein the annealing is done by a mediator, and the negotiating agents are incented to follow a truth-telling strategy that maximizes social welfare.

1 The Challenge: Negotiating Complex Contracts

Work to date on computational models of negotiation has focused almost exclusively on defining contracts consisting of one or a few independent issues [1] [2]. We can frame what these techniques do as follows (see Figure 1 below). Each point on the X axis represents a candidate contract. The Y axis represents the utility of each contract to each agent. Both agents have a reservation utility value: only contracts whose utility is above that agent's reservation value will be accepted. Since relative few issues are involved, the space of all possible contracts can be explored exhaustively, and since the issues are independent, the utility functions mapping a candidate contract to its utility for an agent are *linear* [3], with a single optimum in the utility function for each agent. In such a context, the reasonable strategy is for each agent to start at its own ideal contract, and concede, through iterative proposal exchange, just enough to get the other party to accept the contract. Since the utility functions are simple, it is feasible for one agent to infer enough about the opponent's utility function through observation to make concessions likely to increase the opponent's utility.

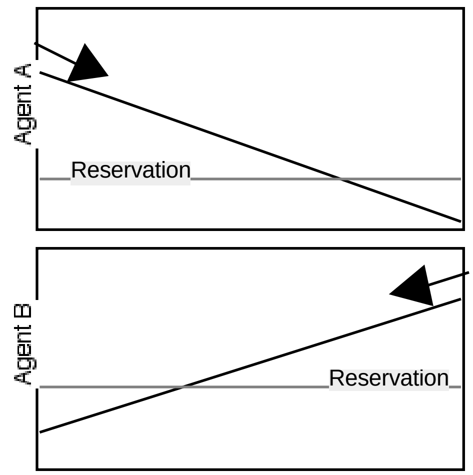

Figure 1: The standard view of negotiation[1].

Real-world contracts, by contrast, are generally much more complex, consisting of a large number of inter-dependent issues. A typical contract may have tens to hundreds of distinct issues. Even with only 50 issues and two alternatives per issue, we encounter a search space of roughly 10^15 possible contracts, too large to be explored exhaustively. The value of one issue selection to an agent, moreover, will often depend on the selection made for another issue. The value to me of a given DVD player, for example, depends on whether it is a good match with the tuner and speakers I plan to purchase with it. Such issue interdependencies lead to *nonlinear* utility functions with multiple local optima [3]. Such contexts, as we have shown in [4], require substantively different negotiation techniques which can allow agents to find 'win-win' contracts in intractably large multi-optima search spaces in a reasonable amount of time.

2 Our Previous Work: Mediated Single Text Negotiation with Annealing Agents

Our efforts to address this challenge have made use of mediated single-text negotiation, a standard approach for dealing with complex negotiations in human settings [5]. In this protocol, a mediator proposes a contract that is either accepted or rejected by each of the parties in the negotiation. A new, hopefully better proposal is then generated by the

---

[1] For simplicity of exposition we show only one dimension in these figures, but there is in actuality one dimension for every issue negotiated over.

mediator based on these responses. This process continues, generating successively better contracts, until the reservation utility value is exceeded for both parties or time runs out. We can visualize this process as follows (Figure 2):

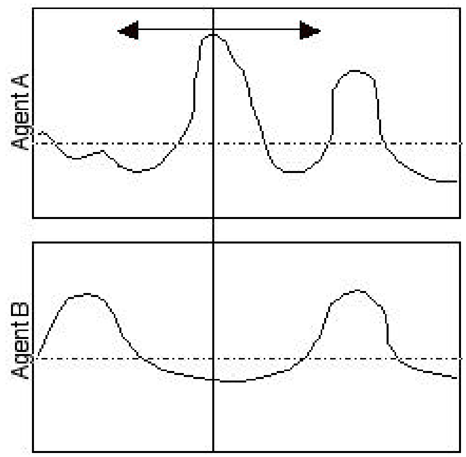

Figure 2: Mediated single text negotiation.

Here, the vertical line represents the contract currently proposed by the mediator. Each new contract moves the line to an adjacent point on the X axis. The goal is to find a contract that is acceptable to both parties. As we can see, if both agents are 'hill-climbers' (i.e. if both agents only accept contracts that are superior to the last mutually accepted one) then mediated single-text negotiation is likely to quickly get stuck in regions where no incremental contract change will increase the utility of the contract to both parties, even though mutually more beneficial contracts exist elsewhere in the contract space.

In an earlier paper [4], we addressed this issue by creating 'annealing' agents that have a time-decreasing willingness to provisionally accept contracts that are inferior to contracts they have already accepted. More specifically, each agent is endowed with a virtual temperature T such that it accepts utility-decreasing contracts with probability:

$$P(\text{accept}) = e^{-\Delta U/T}$$

In other words, the higher the virtual temperature, and the smaller the utility decrement, the greater the probability that the inferior contract will be accepted. The virtual temperature of an annealing agent gradually declines over time so eventually it becomes indistinguishable from a hill-climber. Annealing has proven effective in single-agent optimization, because it can skip utility valleys on the way to higher optima [3]. We hypothesized that annealers would prove superior to hill-climbers in complex contract negotiations.

Negotiations with annealing agents did indeed result in substantially superior final contract utilities, but as the payoff table below shows, there is a catch:

	Agent 2 hill-climbs	Agent2 anneals
Agent 1 hill-climbs	[.86] .73/.74	[.86] .99/.51
Agent1 anneals	[.86] .51/.99	[.98] .84/.84

Table 1: Annealing vs hill-climbing agents.

In this table, the cell values are laid out as follows:

[<social welfare optimality>]
<agent 1 optimality >/<agent 2 optimality >

As we expected, paired hill-climbers do relatively poorly while paired annealers do very well. If we have an annealer paired with a hill-climber, however, the annealer gets 'dragged', by virtue of its initial willingness to concede, into contracts that are beneficial for the hill-climber but very poor for the annealer. While annealing is the socially most beneficial agent strategy, the individually dominant strategy is to hill-climb. Strategic considerations thus drive the agents towards selecting the strategy pairing with the *lowest social welfare.* This is thus an instance of the prisoner's dilemma game. It has been shown that this dilemma can be avoided if we assume repeated interactions between agents [6], but ideally we would prefer to have a negotiation protocol that incents socially beneficial behavior without that difficult-to-enforce constraint.

We were unable, however, to find a way to avoid this dilemma within a single negotiation of this type. If both agents could know ahead of time what strategy the other agent is going to use, then both agents would select annealing. We can not rely on self-reports for this, however, since agents are incented to claim they will use annealing but actually hill-climb. An agent must thus be able to determine the type of its opponent based purely on observing its behavior. It turns out this is relatively easy to do, since annealers accept a much higher percentage of proposed contracts than hill-climber, especially at first. The problem with this approach is that determining the type of an agent based on its voting behavior takes time. By the time the divergence in acceptance rates between annealers and hill-climbers becomes statistically significant, much of the contract utility has already been committed, and it is too late to fully recover from the consequences of having guessed wrong. Another approach we tried is simply to reduce the starting temperature for annealing agents so they are less likely to be dragged away from contracts that are beneficial to them. While selecting a sufficiently low starting temperature does eliminate the annealers' penalty, it results in contracts with much lower final social welfares.

3 The Annealing Mediator

We have developed a negotiation protocol that avoids the prisoner's dilemma in mediated single-text negotiation of complex contracts. The trick is simple: rather than requiring that the negotiating agents anneal, and thereby expose themselves to the risk of being dragged into bad contracts, we moved the annealing into the mediator itself. In our original protocol, the mediator would simply propose modifications of the last contract both negotiating agents accepted. In our refined protocol, the mediator is endowed with a time-decreasing willingness to follow up on contracts that one or both agents rejected. Agents are free to remain hill-climbers and thus avoid the potential of making harmful concessions. The mediator, by virtue of being willing to provisionally pursue utility-decreasing contracts, can skip over valleys in the agents' utility functions and thereby lead the agents to win-win solutions. We describe the details of our protocol, and our evaluations thereof, below.

In our experiments, there were two agents negotiating to find a mutually acceptable contract consisting of a vector S of 40 boolean-valued issues, each issue assigned the value 0 or 1, corresponding to the presence or absence of a given contract clause. This defined a space of 2^40, or roughly 10^12, possible contracts. Each agent had a utility function calculated using its own 40x40 influences matrix H, wherein each cell represents the utility increment or decrement caused by the presence of a given pair of issues, and the total utility of a contract is the sum of the cell values for every issue pair present in the contract:

$$U = \sum_{i=1}^{40} \sum_{j=1}^{40} H_{ij} S_j S_j$$

The diagonal of the influence matrix captures the independent value of a given contract clause, while the other cells capture the dependencies between clause selections. For our experiments, the matrix was initialized to have random values between –1 and +1 in each cell. A different influences matrix was used for each simulation run, in order to ensure our results were not idiosyncratic to a particular configuration of issue inter-dependencies.

The mediator proposes a contract that is initially generated randomly. Each agent then votes to accept or reject the contract. All agents are hill-climbers, which means they accept a mutated contract only if its utility to them is greater than that of the last contract provisionally accepted by the mediator. If both agents vote to accept, the mediator mutates the contract (by randomly flipping one of the issue values) and the process is repeated. If one or both agents vote to reject, the mediator may still follow up on (i.e. propose a mutation of) that contract with a probability given by:

$$P(\text{accept}) = e^{-\Delta U/T}$$

If the mediator does not choose to follow up on the rejected contract, it will propose a mutation of the last contract it decided to accept. The propose-and-vote process is repeated a fixed number of times (in our case 2500) during which time the virtual temperature of the annealing mediator gradually declines to zero. Note that this approach can straightforwardly be extended to N-party (i.e. multi-lateral) negotiations, since we can have any number of parties voting on the contracts.

In our initial implementations each agent gave a simple accept/reject vote for each proposal from the mediator, but we found that this resulted in final social welfares significantly lower than what we earlier achieved using annealing agents. We believe this occurred because the agents provided too little information to effectively guide the mediator's search. In our next round of experiments we accordingly modified the agents so that they provide additional information to the mediator in the form of vote strengths: each agent annotates an accept or reject vote as being *strong* or *weak*. The agents were designed so that there are roughly an equal number of weak and strong votes of each type, so as to maximize the informational content of the vote strength annotations. When the mediator receives these votes, it maps them into numeric values and adds them together according to the following simple scheme:

	Strong accept (1)	**Weak Accept (0)**	**Weak Reject (-1)**	**Strong Reject (-2)**
Strong accept (1)	Accept (2)	Accept (1)	Mixed accept (0)	Weak reject (-1)
Weak Accept (0)	Accept (1)	Accept (0)	Weak reject (-1)	Medium reject (-2)
Weak Reject (-1)	Mixed accept (0)	Weak reject (-1)	Medium reject (-2)	Strong reject (-3)
Strong Reject (-2)	Weak reject (-1)	Medium reject (-2)	Strong reject (-3)	Very strong reject (-4)

Table 2: Scheme for mapping agent votes to overall contract scores.

A proposal is thus accepted by the mediator if both agents voted to accept it, or if it is a 'mixed acccept' (a weak reject by one agent is overridden by a strong accept from the other). The mediator in addition occasionally accepts rejected contracts (i.e. with a negative overall score) using the annealing scheme described above. Note that this scoring scheme allows the system to pursue social welfare-increasing contracts that cause a utility decrement for one agent, which accounts for at least part of its advantage over voting without strength annotations. This approach works very well, achieving final social welfare values that average roughly 99% of optimal[1].

[1] The optimal social welfare was estimated by applying an annealing optimizer to the sum of the two agents' utility functions.

4 Incentives for Truthful Voting

Any voting scheme introduces the potential for strategic non-truthful voting by the agents, and our scheme is no exception. Imagine that one of the agents always votes truthfully, while the other exaggerates so that its votes are always ‘strong’. One might expect that this would bias negotiation outcomes to favor the exaggerator and this is in fact the case:

	Agent 2 exaggerates	Agent 2 tells truth
Agent 1 exaggerates	[.92] .81/.81	[.93] .93/.66
Agent 1 tells truth	[.93] .66/.93	[.99] .84/.84

Table 3: Truth-telling vs exaggerating agents with a simple annealing mediator.

As we can see, even though exaggerating has substantial negative impact on social welfare, agents are individually incented to exaggerate, thus re-creating the prisoner’s dilemma game we encountered in our earlier work. The underlying problem is simple: exaggerating agents are able to induce the mediator to accept all the proposals that are advantageous to them (if they are weakly rejected by the other agent), while preventing the other agent from doing the same. What we need, therefore, is an enhancement to the negotiation protocol that incents truthful voting, thereby preserving equity and maximizing social welfare.

How can this be done? We found that simply placing a limit on the number of strong votes each agent can use does not work. If the limit is too low, we effectively lose the benefit of vote weight information and get the lower social welfares that result. If the strong vote limit is high enough to avoid this, then all an exaggerator has to do is save all of it’s strong votes till the end of the negotiation, at which point it can drag the mediator towards making a series of proposals that are inequitably favorable to it. Similarly, simply limiting the total number of mixed accepts each agent can win for itself during a negotiation is not enough. If the limit is high enough to allow truthful agents to get good social welfares, it is high enough to allow exaggerators to rack up a disproportionate number of mixed accept wins.

The solution, we found, came from enforcing parity between the number of mixed accept wins given to each agent *throughout* the negotiation. If the mediator accepts a proposal that was rejected by one agent and accepted by the other, then the agent that voted to accept accrues one ‘mixed win’. Neither agent is allowed to get more than a given advantage in the mixed win category *at any given point in time*. This way at least rough equity is maintained no matter when (or whether) either agent chooses to exaggerate. The results of this approach were as follows for a win gap limit of 3:

	Agent 2 exaggerates	Agent 2 tells truth
Agent 1 exaggerates	[.91] .79/.79	[.92] .78/.81
Agent 1 tells truth	[.92] .81/.78	[.98] .84/.84

Table 4: Truth-telling vs exaggerating agents with parity-enforcing mediator.

When we have truthful agents, we find that this approach achieves social welfare just slightly below that achieved by a simple annealing mediator, while offering a significantly ($p < 0.01$) higher payoffs for truth-tellers than exaggerators. Truth-telling is thus both the individually dominant and socially most beneficial strategy.

Why does this work? Why, in particular, does a truth-teller fare better than an exaggerator with this kind of mediator? One can think of this procedure as giving agents 'tokens' that they can use to win in mixed vote situations, with the constraint that both agents spend tokens at a roughly equal rate. Recall that in this case a truthful agent, offering a mix of strong and weak votes, is paired with an exaggerator for whom all of the accepts and rejects are strong ones. The truthful agent can therefore only win a mixed vote via annealing (see Table 2), and this is much more likely when its vote was a strong accept rather than a weak one. In other words, the truthful agent spends its tokens predominantly on contracts that truly offer it a strong utility increase. The exaggerator, on the other hand, spends its tokens indiscriminately, trying to elicit a mixed win even when the utility increment it derives is relatively small. At the end of the day, the truthful agent has spend its tokens more wisely and to better effect.

5 Contributions

The key contribution of this paper is to describe and evaluate a mediated single text negotiation protocol that is well-suited for non-linear negotiations and avoids an incentive structure that makes globally unhelpful behavior individually rational. The approach involves two key ideas: (1) using a mediator that selectively ignores reject votes in order to skip 'valleys' in the agents' multi-optima utility functions, and (2) using a parity enforcement mechanism to ensure that truth-telling is individually rational. We believe that these insights may prove useful in a range of interdependent synthesis tasks (such as collaborative design, scheduling and planning).

6 Next Steps

The high social welfare values achieved by our approach partially reflect the fact that the utility functions for each agent, based as they are solely on binary dependencies, are relatively easy to optimize. Higher-order dependencies, common in many contexts, are known to generate more challenging utility landscapes [7] and will be addressed in future work. In our experiments the contract space was explored in random walk fashion, and all the ‘intelligence’ was in the evaluation process. More focused contract generation operators may prove important. Finally, it is clear that if agents cooperate they can produce higher contract utilities. We will explore whether additional truthful information revelation may prove rational for agents as a way of making sure that complex mutually agreeable contracts can be found in an acceptable amount of time.

7 Acknowledgements

This work was supported by funding from the DARPA Control of Agent-Based Systems (CoABS) program, the NSF Computation and Social Systems program, and Neptune Inc.

8 References

1. Faratin, P., C. Sierra, and N.R. Jennings, *Using similarity criteria to make negotiation trade-offs.* Proceedings Fourth International Conference on MultiAgent Systems. IEEE Comput. Soc., 2000.
2. Ehtamo, H., E. Ketteunen, and R. Hamalainen, *Searching for Joint Gains in Multi-Party Negotiations.* European Journal of Operational Research, 2001. **1**(30): p. 54-69.
3. Bar-Yam, Y., *Dynamics of complex systems*. 1997, Reading, Mass.: Addison-Wesley. xvi, 848.
4. Klein, M., P. Faratin, H. Sayama, and Y. Bar-Yam. *Negotiating Complex Contracts*. in *Autonomous Agents and Multi-Agent Systems*. 2002. Bologna Italy: AAAI Press.
5. Raiffa, H., *The art and science of negotiation*. 1982, Cambridge, Mass.: Belknap Press of Harvard University Press. x, 373.
6. Axelrod, R., *The Evolution Of Cooperation*. 1984: Basic Books.
7. Kauffman, S.A., *The origins of order: self-organization and selection in evolution*. 1993: Oxford University Press.