Negotiations Over Time in a Multi-Agent Environment Preliminary Report

Sarit Kraus*
Graduate School for Library Studies and
Dept. of Computer Science
Hebrew University, Jerusalem 91904 Israel
sarit@cs.huji.ac.il

Jonathan Wilkenfeld
Dept. of Government and Politics
University of Maryland
College Park, MD 20742
wilkenfeld@umd2.umd.edu

Abstract

One of the major foci of research in distributed artificial intelligence (DAI) is the design of automated agents which can interact effectively in order to cooperate in problem-solving. Negotiation is recognized as an important means by which inter-agent cooperation is achieved. In this paper we suggest a strategic model of negotiation for N agents ($N > 3$), that takes the passage of time during the negotiation process itself into consideration. Changes in the agent's preferences over time will change their strategies in the negotiation and, as a result, the agreements they are willing to reach. We will show that in this model the delay in reaching such agreements can be shortened and in some cases avoided altogether.

1 Introduction

One of the major foci of research in distributed artificial intelligence (DAI) is the design of automated agents which can interact effectively in order to cooperate in problem-solving. Negotiation is recognized as an important means by which inter-agent cooperation is achieved. That is, DAI is concerned with the design of agents which are able to communicate in such a way as to enhance the possibility of reaching mutually beneficial agreements concerning problems such as a division of labor or resources among the agents.

Negotiation has always been a central theme in DAI research [Davis and Smith, 1983; Georgeff, 1983; Malone *et* al., 1988; Durfee, 1988; Durfee and Lesser, 1989; Rosenschein and Genesereth, 1985; Sathi and Fox, 1989; Conry *et* a/., 1988; Zlotkin and Rosenschein, 1990]. This research has focused on strategies for designing agents capable of reaching mutually beneficial agreements. Sycara ([Sycara, 1987]), using case-based reasoning, and Kraus et al. ([Kraus *et* al., 1991J) modeled negotiations from a cognitive standpoint.

*This work was partially completed while the first author was at the Institute for Advanced Computer Studies and Dept. of Computer Science, University of Maryland, College Park.

Yet it is also recognized that although negotiations are necessary for reaching such agreements, the negotiation process is both costly and time-consuming, and thus may increase the overhead of the operation in question (see [Bond and Gasser, 1988]). In negotiations on such issues as job-sharing or resource allocation, it is important to minimize the amount of time spent on negotiating mutually beneficial agreements so as not to detract from time spent on the task itself. Thus, in the presence of time constraints, negotiation time should be taken into consideration.

In [Kraus and Wilkenfeld, 1991a] we propose a strategic model of negotiation that takes the passage of time during the negotiation process itself into consideration. That study focused exclusively on a two-agent model. The present study generalizes this process by considering the N-agent environment.

Following [Rosenschein and Genesereth, 1985; Zlotkin and Rosenschein, 1990; Kraus and Wilkenfeld, 1990; Kraus and Wilkenfeld, 1991a] we examine negotiation using game theory techniques with appropriate modifications to fit artificial intelligence situations. We will focus primarily on works in game theory and economics that have studied the effect of time preferences on the negotiation process, following the classic paper by Rubinstein ([Rubinstein, 1982]). Unlike the work of Zlotkin and Rosenchein, [Zlotkin and Rosenschein, 1990] we investigate multi-agent environments (more than two agents) and our approach makes no assumptions about the protocol for negotiations. Also, our model takes the passage of time during the negotiation process itself into consideration. Furthermore, by taking the passage of time during the negotiation process into consideration, our approach is able to influence the outcome of the negotiation so as to avoid delays in reaching agreements.

2 Initial Setting

N autonomous agents $A\backslash,\ A_2,...,As$ have a common goal they want to satisfy as soon as possible. In order to satisfy a goal, costly actions must be taken and an agent cannot satisfy the goal without reaching an agreement with one of the other agents. Each of the agents wants to minimize its costs, i.e., prefers to do as little as possible, if it can assume that the goal will be fulfilled properly without additional effort on its part. We note that even though the agents have the same goal (under our simplified assumptions), there is actually a conflict of interests. The agents have different preferences concerning goal satisfaction and for the different possible agreements which can be reached.

We make the following assumptions:

1. Full information - each agent knows all relevant information including the other agents' preferences for the different outcomes over time.
2. The agents are rational - they will behave according to their preferences.
3. Commitments are enforced - if an agreement is reached both sides are forced to follow it.
4. Assumptions (l)-(3) are common knowledge.

We demonstrate the cases we are interested in with the following example.

Example 2.1 *Three robots, A, B and C, stationed on a satellite, are instructed to move an expensive telescope from one location to another as soon as possible. Delay in moving the telescope will reduce the number of pictures sent back to scientists on earth. Any two of the robots can move the telescope, but the tools essential to perform the task are distributed among the three agents. Any of the agents can opt out of the negotiation, choosing not to satisfy the goal. If that occurs, the remaining two cannot achieve the goal (since critical tools will be missing).*

3 The Structure of Negotiations

Our strategic model of negotiations is a model of Alternative Offers.[1]

For reasons of simplification and clarity, we will concentrate on the case of three agents, but our results can be easily extended to N agents where an agent can satisfy a goal by reaching an agreement with another agent[2]. So, in our case, three agents, A, B and C have a common goal they want to satisfy as soon as possible. Each of them has a set of capabilities, P_A, P_B and P_C and a set of tools T_A, T_B and T_C respectively. The agents' capabilities influence their ability and their ways of satisfying the goal, and all the tools are needed to satisfy the goal. We now present formal definitions of agreements and strategies.

3.1 Agreements and Strategies

We first define the set of possible agreements. We assume that there exists a set of possible agreements between any two agents. We denote by $\mathcal{A}$ the set of agents.

Definition 3.1 Agreement: *Let Act be the set of actions required to satisfy a goal. An agreement is a pair (s_i, s_j), where $i, j \in \mathcal{A}$, $i \neq j$; $s_i \cup s_j = Act$. s_i is agent i's portion of the work*[3]. *We assume that the set S_{ij} $i, j \in \mathcal{A}, i \neq j$ includes all the possible agreements between agents i and j. We also assume that $S_{ij} = S_{ji}$. Let $S \stackrel{\text{def}}{=} \cup_{ij} S_{ij}$ $i, j \in \mathcal{A}$ $i \neq j$, i.e., S is the set of all possible agreements.*

The negotiation procedure is as follows. The agents can take actions only at certain times in the set $\mathcal{T} = \{0, 1, 2...\}$. In each period $t \in \mathcal{T}$ one of the agents, say i, proposes an agreement to one of the other agents. The other agent (j) either accepts the offer (chooses Y) or rejects it (chooses N), or opts out of the negotiation (chooses O). Also, the third agent may opt out of the negotiation (chooses O), or it can choose not to do anything (chooses Nop). If the offer is accepted, without the third agent opting out, then the negotiation ends and the agreement is implemented (i.e., each of the agents that reached the agreement, does its part of the job and the other agent is obliged to contribute its tools). Also, opting out by one of the agents ends the negotiations since all the tools are required to satisfy the goal. After a rejection, another agent must make a counter offer, and so on.

There are no rules which bind the agents to any previous offers and there is no limit on the number of periods. The only requirement we make is that the length of a single period be fixed and the agents always make offers in the same order, i.e., agent i makes offers in time periods t, $t + 3$, $t + 6...$, agent j makes offers on periods $t + 1$, $t + 4...$, and similarly to the last agent. If an agreement is never reached, and none of the agents opts out, we assume the outcome to be D (Disagreement). We do not make any assumption about who begins the negotiations, i.e., who makes the first offer and who is the second agent to make an offer. So, without loss of generality we assume that $i, j, l \in \mathcal{A}, i \neq j \neq l$ are the first, the second and the third agents, respectively.

Definition 3.2 Negotiation Strategies: *A strategy is a sequence of functions. The domain of the ith element of a strategy is a sequence of agreements of length i and its range is the set $\{Y, N, O, Nop\} \cup S$. We first define a strategy f for an agent i who is the first agent to give an offer.*

Let $f = \{f^t\}_{t=0}^{\infty}$, where $f^0 \in \{S_{ij} \cup S_{il}\}$, for $t = 3n, n \in \mathcal{T}$ $f^t : S^t \rightarrow \{S_{ij} \cup S_{il}\}$, and for $t = 3n + 1$, $n \in \mathcal{T}$ $f^t : S^t \times \{S_{ji} \cup S_{jl}\} \rightarrow \{Y, N, O, Nop\}$ where if $s^{t+1} \in S_{ji}$, $f^t(s^0, s^1, ..., s^{t+1}) \in \{Y, N, O\}$, and if $s^{t+1} \in S_{jl}$, $f^t(s^0, s^1, ..., s^{t+1}) \in \{O, Nop\}$, ($S^t$ is the set of all sequences of length t of elements in S). For $t = 3n + 2$, $n \in \mathcal{T}$ $f^t : S^t \times \{S_{li} \cup S_{lj}\} \rightarrow \{Y, N, O, Nop\}$ where if $s^{t+1} \in S_{li}$, $f^t(s^0, s^1, ..., s^{t+1}) \in \{Y, N, O\}$, and if $s^{t+1} \in S_{lj}$, $f^t(s^0, s^1, ..., s^{t+1}) \in \{O, Nop\}$. We denote by F the set of all strategies of the agent who starts the bargaining.

Similarly, we denote by G the set of all strategies of the agent j who is the second to make offers and we denote by H the set of all strategies of agent l who is the third

[1]See [Osborne and Rubinstein, 1990] for a detailed review of the bargaining game of Alternating Offers.

[2]In the case in which the agents may divide the labor between all of them, i.e., agreement may be reached only among all the agents, the model of Alternative Offers is usually disappointing (see [Osborne and Rubinstein, 1990]). It is still useful when the agents have specific types of utility functions (see [Stahl, 1977]).

[3]A similar definition can be given concerning a division of resources.

to make offers[4].

Let $\sigma(f,g,h)$ be a sequence of offers possibly ending with O in which agent i (who can be either A, B or C) starts the bargaining and adopts $f \in F$, agent j adopts $g \in G$ and agent l adopts $h \in H$. Let $T(f,g,h)$ be the length of $\sigma(f,g,h)$ (may be ∞). Let $La(f,g,h)$ be the last element of $\sigma(f,g,h)$ (if there is such an element). $La(f,g,h)$ may be either in S or O which denotes that one of the agents opted out without the agents reaching an agreement.

Definition 3.3 Outcome of the negotiation: *The outcome function of the game is defined by*

$$P(f,g,h) = \begin{cases} D & \text{if } T(f,g,h) = \infty \\ (La(f,g,h), T(f,g,h) - 1) & \text{otherwise} \end{cases}$$

Thus, the outcome (s,t) where $s \in S$ is interpreted as the reaching of an agreement s in period t. The agreement can be between any two agents. (O,t) is interpreted as one of the agents opting out of the negotiations, and the symbol D indicates a perpetual disagreement without any agent opting out.

3.2 Agents' Preferences Over Possible Outcomes

The last component of the model is the preferences of the agents on the set of outcomes. Each agent has preferences for agreements reached at various points in time, and for opting out at various points in time. The time preferences and the preferences between agreements and opting out are the driving force of the model.

Formally, we assume that any agent $i \in \mathcal{A}$ has a preference relation $\succeq_i$ on the set $\{S \times \mathcal{T}\} \cup \{\{O\} \times \mathcal{T}\} \cup \{D\}$, where $\mathcal{T}$ is the set of time periods.

We note here that by defining an outcome to be either a pair (s,t) or (O,t) or D, we have made a restrictive assumption about the agent's preferences. We assume that agents care only about the nature of the agreement or opting out, and the time at which the outcome is reached, and not about the sequence of offers and counteroffers that leads to the agreement. In particular, no agent regrets either having made an offer that was rejected or rejecting an offer (see, for example, the discussion of "decision-regret" in [Raiffa, 1982]).

We make some assumptions about the agents preferences. First we assume that the least-preferred outcome is disagreement (D).

(A0) For every $s \in S$ and $t \in \mathcal{T}$, $(s,t) \succ_i D$ and $(O,t) \succ_i D$ (Disagreement is the worst outcome).

The next two conditions (A1), (A2) concern the behavior of $\succ_i$ on $S_{ij} \times \mathcal{T}$, i.e., concerning agreements reached with another agent in different time periods. We will assume that the agents have no preference among the actions in Act , i.e., all actions are equally difficult. Condition (A1) requires that among agreements reached in the same period with the same agent, agent i prefers fewer numbers of actions s_i.

(A1) For $i \in \mathcal{A}$, $t \in \mathcal{T}$ and $s,r \in S_{ij}$, $j \in \mathcal{A}, j \neq i$, if $|r_i| < |s_i|$, then $(r,t) \succ_i (s,t)$.

[4]The full definitions of the j and l's strategies can be found in [Kraus and Wilkenfeld, 1991b].

We note that this condition does not hold for comparisons among agreements between different agents, i.e., beside the amount of work the agents should do it has other considerations that depend on the other agent's capacity to do its part of the job (P_j).

The next assumption greatly simplifies the structure of preferences among the agreements between any two agents. It requires that preferences between (s_1,t_1) and (s_2,t_2) where $s_i \in S_{ij}$, depend only on s_1, s_2 and the differences between t_1 and t_2.

(A2) For all $r,s \in S_{ij}$, $t_1,t_2,\delta \in \mathcal{T}$ and $i \in \mathcal{A}$, $(r,t_1) \succeq_i (s,t_1+\delta)$ iff $(r,t_2) \succeq_i (s,t_2+\delta)$ (Stationarity).

We note that assumption (A2) does not hold for O and for comparisons among agreements between different agents[5].

Example 3.1 *In the situation from Example 2.1, agents B and C believe that the safest way to move the telescope is by joint action of A and C. In such a case the only possible agreement is one in which agent A will do most of the work. On the other hand agent A prefers to reach an agreement with agent B. In such a case it will need to take fewer actions, and it prefers this despite the cost in safety of the move. It is not safe at all for agents B and C to move the telescope and none of the agents prefers such an agreement.*

When analyzing the model, the main question is whether a possibility exists that the agents will reach an agreement. An important feature of the model that strongly influences the outcome of the game is the preference of a player between an agreement and opting out. We need the following definition in order to compare between agreements and opting out.

Definition 3.4 *For every $t \in \mathcal{T}$ and $i \in \mathcal{A}$ we define $\bar{S^i}^t_{jl} = \{s | s \in S_{jl}, (s,t) \succ_i (O,t), i,j,l \in \mathcal{A}, j \neq l\}$. Let $\bar{S^i}^t = \bar{S^i}^t_{ij} \cup \bar{S^i}^t_{il} \cup \bar{S^i}^t_{jl}$ where $i,j,l \in \mathcal{A}, i \neq j \neq l$. If $\bar{S^i}^t \neq \emptyset$, we denote $\hat{s}^{i,t} = \min_{\prec_i} \bar{S^i}^t$, i.e., $\hat{s}^{i,t}$ is the worst agreement that can be reached in period t which is still better to agent i than opting out. In case such an agreement does not exist, we define $\hat{s}^{i,t} = -1$.*

We would like now to introduce an additional assumption that will ensure that if all agents prefer some agreements over opting out, an agreement can be reached.

(A3) For any $t \in \mathcal{T}$, if for every $i \in \mathcal{A}$, $\hat{s}^{i,t} \neq -1$, then $\bar{S^i}^t \cap \bar{S^j}^t \cap \bar{S^l}^t \neq \emptyset$, where $i,j,l \in \mathcal{A}, i \neq j \neq l$.

3.3 Perfect Equilibrium

The main question is how a rational agent chooses his strategy for the negotiation. A useful notion is the Nash Equlibrium ([Nash, 1950; Luce and Raiffa, 1957]). A triplets of strategies (σ,τ,ϕ) is a Nash Equilibrium if, given τ and ϕ, no strategy of agent A can result in an outcome that agent A prefers to the outcome generated by (σ,τ,ϕ) and, similarly, to agent B given σ and ϕ and to agent C given σ and τ (assuming that A, B and C

[5]An example of a utility function which satisfies the above assumptions concerning agreements is the following: $U_i(s,t) = C_j + |Act - s_i| + t * c_i$ where $s \in S_{ij}$.

are the first, second and third agents to make an offer, respectively). If there is a unique equilibrium, and if it is known that an agent is designed to use this strategy, no agent will prefer to use a strategy other than these. However, the use of Nash Equilibrium is not an effective way of analyzing the outcomes of the models of Alternating Offers since it puts few restrictions on the outcome and yields too many equilibria points (see the proof in [Rubinstein, 1982]). Therefore, we will use the stronger notion of **(subgame) perfect equilibrium** (P.E.) (see [Selten, 1975]) which requires that the agents' strategies induce an equilibrium in any stage of the negotiation, i.e., in each stage of the negotiation, assuming that an agent follows the P.E. strategy, the other agent does not have a better strategy then to follow its own P.E. strategy. So, if there is a unique perfect equilibrium, and if it is known that an agent is designed to use this strategy, no agent will prefer to use a strategy other than this one in each stage of the negotiation.

4 All Agents Lose Over Time

Suppose all agents are losing over time. We assume that all agents prefer to reach a given agreement sooner rather than later, and that all agents prefer to opt out sooner rather than later. We also assume that if a player prefers an agreement over opting out in some period t, then it prefers the same agreement in time period t' prior to t over opting out in t'. Formally:

(A4) For any $i \in \mathcal{A}$ and $t, t_1, t_2 \in \mathcal{T}$, if $t_1 < t_2$ then $(O, t_1) \succ_i (O, t_2)$, and for any $s \in S$, $(s, t_1) \succ_i (s, t_2)$. If $(s,t) \succ_i (O,t)$, then for any $t' \in \mathcal{T}$ such that $t' < t$, $(s,t') \succ_i (O,t')$. If $(s,t) \succ_i (O,t+1)$, then for any $t' \in \mathcal{T}$ such that $t' < t$, $(s,t') \succ_i (O,t'+1)$.

We also assume that all agents prefer to take part in satisfying the goal. Formally:

(A5) For any agent $i \in \mathcal{A}$, $t \in \mathcal{T}$ and for agreements $s \in \{S_{ij} \cup S_{il}\}$ and $s' \in S_{jl}$, $j, l \in \mathcal{A}, j \neq l \neq i$, $(s,t) \succ_i (s',t)$.

The first case we consider is that there is a period in which one of the agents prefers opting out of the negotiation over any agreement. We also assume that in such a case the other agents still prefer at least one agreement over opting out in this period. We will now prove that in such a case if the game has not ended in prior periods, then an agreement will be reached in the prior period, or two periods prior, to the period in which one of the agents prefers opting out over any agreement. We will use additional notation. $\hat{S}^j_t = \{s | s \in S_{jl} \cup S_{ji}, (s,t) \succ_j (O, t+1)\}$ where $t \in \mathcal{T}, i, j, l \in \mathcal{A}, i \neq j \neq l$. $\hat{S}^j_t$ includes the offers agent j can make on time period t which are better than opting out in the next time period.

Lemma 1 *Let $(\hat{f}, \hat{g}, \hat{h})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A5. Suppose there exists $T \in \mathcal{T}$ such that $T > 1$ and for any $t \in \mathcal{T}$, and for any $i \in \mathcal{A}$ if $t < T$ then $\hat{s}^{i,t} \neq -1$ and there exists $i \in \mathcal{A}$ such that $\hat{s}^{i,T+1} = -1$. Suppose it is j's turn, $j \in \mathcal{A}$, to make an offer in time period T. If $\hat{S}^j_T \cap \bar{S^l}^T \cap \bar{S^i}^T \neq \emptyset$, where $i, l \in \mathcal{A}, i \neq j \neq l$, then let $s^j = max_{\prec_j}\{\hat{S}^j_T \cap \bar{S^l}^T \cap \bar{S^i}^T\}$. Let $s^j \in S_{jk}$, $k = i$ or $k = l$, then using its perfect equilibrium strategy, j will offer k s^j and k will accept the offer.*

If $\hat{S}^j_T \cap \bar{S^l}^T \cap \bar{S^i}^T = \emptyset$ then suppose, without loss of generality, that it is i's turn to give an offer in time period $T-1$ then let $s^i = max_{\prec_i}\hat{S}^i_{T-1} \cap \bar{S^l}^{T-1} \cap \bar{S^j}^{T-1}$. Let $s^i \in S_{ik}$, $k = j$ or $k = l$, then using its perfect equilibrium strategy, i will offer k, in time period $T-1$, s^i and k will accept the offer.

In the first case we denote s^j by $\hat{s}$ and T by $\hat{T}$ and in the second case we denote s^i by $\hat{s}$ and $T-1$ by $\hat{T}$.

Proof: The proof of this lemma and the following lemmas and theorems can be found in [Kraus and Wilkenfeld, 1991b].

We will show that under certain conditions, in any period there will be an agreement that will be accepted.

Lemma 2 *Let $(\hat{f}, \hat{g}, \hat{h})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A5 and let $\hat{s}$ and $\hat{T}$ be defined as in Lemma 1. Suppose $T < \hat{T}$ such that it is i's turn to make an offer. If for $j \in \mathcal{A}, j \neq i$ there exists $x^T_{ij} \in S_{ij}$ satisfies the following conditions:*

1. *$(x^T_{ij}, T) \succ_k (O,T)$ where $k \in \mathcal{A}$ and $i \neq k$, and $(x^T_{ij}, T) \succ_i (O, T+1)$.*
2. *$(x^T_{ij}, T) \succ_k (\hat{s}, \hat{T})$, where $k = i$ or $k = j$.*
3. *If exist $t \in \mathcal{T}, T < t < \hat{T}$ and $s \in S$, such that $(s,t) \succ_j (x^T_{ij}, T)$ then $(O, T+1) \succ_k (s,t)$ where $k \in \mathcal{A}$, $k \neq j$.*

then if agent i offers agent j x^T_{ij} in period T, then j using its P.E. strategy, will accept the offer. We denote the set of all the x^T_{ij}'s satisfying the above conditions by X^T_{ij}.

We will show now that in some cases, the negotiations are concluded in the first period. In other cases, some delay in reaching an agreement may occur. In the worst case, agreement will be reached only in period $\hat{T}$.

Theorem 1 *Let $(\hat{f}, \hat{g}, \hat{h})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A5 and let X^T_{ij} be defined as in Lemma 2. Suppose for any $i, j \in \mathcal{A}, i \neq j$, $X^{\hat{T}-1}_{ij} \neq \emptyset$. If for all X^t_{ij} such that $i, j \in \mathcal{A}$, $i \neq j$, $t \in \mathcal{T}$ and $t < \hat{T}$, if $x \in X^t_{ij}$ then if there is $s \in S$, such that $(s,t) \succ_j (x, t-1)$ then $(O,t) \succ_k (s,t)$, for any $k \in \mathcal{A}, k \neq j$, then when the agents use their P.E. strategies, agreement will be reached in the first, second or the third periods.*

5 Time is Valuable Only to Some Agents

Suppose one of the agents does not lose over time and even gains at least in the early stages of the negotiation. For example, in the robots in Example 2.1, suppose agent A controls the telescope and uses it for his purposes until an agreement is reached. In this case we will assume that agent A prefers any agreement over opting out. As in example 3.1 we will assume that the set of agreements between B and C is empty and that there is only one possible agreement between C and A. Furthermore, we assume that both B and C prefer the agreement between

A and C over any agreement between A and B[6]. In this case we show that even when one of the agents gains over time but prefers not to opt out, if the other two agents, which lose over time, have the same preferences they can force the first agent to meet their conditions.

Formally we will make the following assumptions.

(A6) $S_{CB} = \{\emptyset\}$[7], and $S_{AC} = \{s_{AC}\}$. (The size of the set of possible agreements is limited).

(A7) For every $t \in T$ and for every $s \in S_{AB}$, $(s_{AC},t) \succ_i (s,t)$, $i \in \{B,C\}$ and $(s,t) \succ_A (s_{AC},t)$. (Agents B and C have contradictory preferences to agent A).

(A8) For any $t_1,t_2 \in T$ such that $t_1 < t_2$ and for any $s \in S$, $(s,t_2) \succ_A (s,t_1)$, $(O,t_2) \succ_A (O,t_1)$ and for $i \in \{B,C\}$, $(s,t_1) \succ_i (s,t_2)$, $(O,t_1) \succ_i (O,t_2)$. (Agent A gains over time and agents B and C lose over time).

(A9) For any $t \in T$ and for any $s \in S$, $(s,t) \succ_A (O,t)$. (Agent A prefers agreements over opting out).

(A10) If for any $t,t' \in T$ and for any $s \in S_{AC} \cup S_{AB}$, if $(s,t) \succ_B (O,t')$ then $(s,t) \succ_C (O,t')$ and if $(s,t) \succ_i (O,t)$, where $i \in B,C$ and $t' < t$ then $(s,t') \succ_i (O,t')$.

We will show now that if there is a period where agent B prefer opting out over any agreement then negotiations will be concluded before that period even though agent A gains over time.

Lemma 3 *Let $(\hat{f},\hat{g},\hat{h})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A2 and A6-A10. If there exist $t_1,t_2 \in T$ such that $t_1 < t_2$, for any $s \in S_{AB}$, $(O,t_1) \succ_B (s,t_1)$, $(s_{AC},t_2-1) \succ_B (O,t_2-1)$ and $(O,t_2) \succ_B (s_{AC},t_2)$, then $P(\hat{f}^{t_1},\hat{g}^{t_1},\hat{h}^{t_1}) = (s_{AC},\hat{t})$[8] $\hat{t} \in T$ where $\hat{t}$ is the maximal t such that it is either A or C's turn to make an offer, $t_1 \le \hat{t} < t_2$ and $(s_{AC},\hat{t}) \succ_B (O,t_1)$.*

The next theorem describes the behavior of the agents.

Theorem 2 *Let $(\hat{f},\hat{g},\hat{h})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A2 and A6-A10. If there exist $t_1,t_2 \in T$ such that $t_1 < t_2$, for any $s \in S_{AB}$, $(O,t_1) \succ_B (s,t_1)$, $(s_{AC},t_2-1) \succ_B (O,t_2-1)$ and $(O,t_2) \succ_B (s_{AC},t_2)$, then if there is no $t < \hat{t}$ ($\hat{t}$ is defined in lemma 3) such that $(O,t) \succ_B (s_{AC},\hat{t})$ and there is no $t < \hat{t}$ and $s \in S_{AB}$ such that $(s,t) \succ_B (s_{AC},\hat{t})$ and $(s,t) \succ_A (s_{AC},\hat{t})$ then $P(\hat{f},\hat{g},\hat{h}) = (s_{AC},\hat{t})$.*

We will demonstrate the above results with the robots from Examples 2.1 and 3.1.

Example 5.1 *In the situation in examples 2.1 and 3.1, suppose the robots satisfy A0-A2 and A6-A10 and suppose $\forall s \in S_{AB}$, $(O,2) \succ_B (s,2)$ and $\forall t \in T, t < 2$ there exists $s^t \in S_{AB}$ such that $(s^t,t) \succ_B (O,t)$, and $(O,10) \succ_B (s_{AC},10)$ and $(s_{AC},9) \succ_B (O,9)$. Suppose also that $(s_{AC},3) \succ_B (O,0)$ and for any $s \in S_{AB}$, $(s_{AC},3) \succ_B (s,0)$ but $\forall t \in T, t > 3, (O,2) \succ_B (s_{AC},t)$. Suppose, A, B and C are the first, second and third agents to make an offer, respectively. In period 3 A will offer C to move the telescope together and C will agree.*

We will examine now the case where there is a time period smaller than $\hat{t}$ and an agreement between A and B that both agents prefer to reach in this time period over s_{AC} in period $\hat{t}$. We will need some additional notations for dealing with this case.

Definition 5.1 *Let $\hat{X} \stackrel{\text{def}}{=} \{(s,t) | s \in S_{AB}, (s,t) \succ_i (s_{AC},\hat{t})$, where $i \in \{A,B\}, t \in T$ and $t < \hat{t}\}$ ($\hat{t}$ is defined in lemma 3). Let $\hat{x}^i \stackrel{\text{def}}{=} max_{\succ_i}\hat{X}, i \in \{A,B\}$.*

Theorem 3 *Let $(\hat{f},\hat{g},\hat{h})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A2 and A6-A10. If there exist $t_1,t_2 \in T$ such that $t_1 < t_2$, for any $s \in S_{AB}$, $(O,t_1) \succ_B (s,t_1)$, $(s_{AC},t_2-1) \succ_B (O,t_2-1)$ and $(O,t_2) \succ_B (s_{AC},t_2)$, then if there is no $t < \hat{t}$ ($\hat{t}$ is defined in lemma 3), such that $(O,t) \succ_B (s_{AC},\hat{t})$ and if $\hat{x}^A = (s',t') = \hat{x}^B$ then $P(\hat{f},\hat{g},\hat{h}) = (s',t')$.*

6 The Application of the Theory in Building Autonomous Agents

One of the main questions is how one can use the above theoretical results in building agents capable of acting and negotiating under time constraints.

We note that in each of the cases we have investigated, either in this paper or in [Kraus and Wilkenfeld, 1991a], where we presented a strategic model of negotiations for only two agents, the perfect-equilibrium strategies are determined by parameters of the situation.

So, one can supply agents with the appropriate strategies for each of the cases we have dealt with. When an agent participates in one of those situations, it will need to recognize which type of situation it is in. Assuming the agent is given the appropriate arguments about the situation it is involved in it can construct the exact strategy for its specific case and use it in the negotiations. Since we provide the agents with unique perfect equilibrium strategies, if we announce it to the other agents in the environment, the other agents can not do better than to use their similar strategies.

7 Conclusion and Future Work

In this paper we demonstrate how the incorporation of time into the negotiation procedure contributes to a more efficient negotiation process where there are at least three agents in the environment. We show that in different cases this model, together with the assumption that the agents' strategies induce an equilibrium in any stage of the negotiation, may result in the agent being able to use negotiation strategies that will end the negotiation with only a small delay. We suggest that these results are useful in particular in situations with time constraints. We are in the process of using this model in developing agents that will participate in crisis situations where time is an important issue.

The most obvious outstanding question concerns the relaxation of the assumption of complete information.

---

[6] Other examples where such situations occurred can be found in [Kraus and Wilkenfeld, 1991a; Kraus and Wilkenfeld, 1990].

[7] $\emptyset$ denotes the empty agreement.

[8] Let $s^0 ... s^T \in S$ define $f|s^0...s^T$ to be the strategy dervied from f after the offers $s^0,...,s^T$ have been announced and already rejected. If $f^T(s_0,...,s_{T-1})$ does not depend on $s_0,...s_{T-1}$ we denote the result by f^T. Similarly for g and h.

In many situations the agents do not have full information concerning the other agents. Several works in game theory and economics have considered different versions of the model of Alternative Offers with incomplete information (see for example, [Rubinstein, 1985; Osborne and Rubinstein, 1990; Chatterjee and Samuelson, 1987]). We are in the process of modifying those results for use in DAI environments.

8 Acknowledgement

We would like to thank K. Arrow, R. Aumann, B. Bueno de Mesquita, J. Oppenheimer, A. Rubinstein and P. Young for helpful suggestions about game theory and negotiation. We would also like to thank J. Hendler and D. Perlis for helpful discussions.

References

[Bond and Gasser, 1988] A. H. Bond and L. Gasser. An analysis of problems and research in DAL In A. H. Bend and L. Gasser, editors, *Readings in Distributed Artificial Intelligence,* pages 3-35. Morgan Kaufmann Publishers, Inc., San Mateo, California, 1988.

[Chatterjee and Samuelson, 1987] K. Chatterjee and L. Samuelson. Bargaining with two-sided incomplete information: An infinite horizon model with alternating offers. *Review of Economic Studies,* 54:175-192, 1987.

[Conry *et al,* 1988] S. E. Conry, R. A. Meyer, and V. R. Lesser. Multistage negotiation in distributed planning. In A. H. Bond and L. Gasser, editors, *Readings in Distributed Artificial Intelligence,* pages 367-384. Morgan Kaufmann Publishers, Inc., San Mateo, California, 1988.

[Davis and Smith, 1983] R. Davis and R.G. Smith. Negotiation as a metaphor for distributed problem solving. *Artificial Intelligence,* 20:63-109, 1983.

[Durfee and Lesser, 1989] E. H. Durfee and Victor R. Lesser. Negotiating Task Decomposition and Allocation Using Partial Global Planning. In L. Gasser and M.N. Huhns, editors, *Distributed Artificial Intelligence, Volume II,* pages 229-244. Pitman/Morgan Kaufmann, London, 1989.

[Durfee, 1988] E. H. Durfee. *Coordination of Distributed Problem Solvers.* Kluwer Academic Publishers, Boston, 1988.

[Georgeff, 1983] M. Georgeff. Communication and interaction in multi-agent planning. In *Proc. of the National Conference on Artificial Intelligence,* pages 125-129, Washington, D.C., 1983.

[Kraus and Wilkenfeld, 1990] S. Kraus and J. Wilkenfeld. An automated strategic model of negotiation. In *AAAI-90 Workshop on Reasoning in Adversarial Domains,* Boston, 1990.

[Kraus and Wilkenfeld, 1991a] S. Kraus and J. Wilkenfeld. The function of time in cooperative negotiations. In *Proc. of AAAI-91,* California, 1991. To appear.

[Kraus and Wilkenfeld, 1991b] S. Kraus and J. Wilkenfeld. Negotiations over time in a multi agent environment. Technical Report UMIACS TR 91-51 CS TR 2649, Institute for Advanced Computer Studies, University of Maryland, 1991.

[Kraus *et* al., 1991] S. Kraus, E. Ephrati, and D. Lehrnann. Negotiation in a non-cooperative environment. *Journal of Experimental and Theoretical Artificial Intelligence,* 1991. Accepted for publication.

[Luce and RaifTa, 1957] R. D. Luce and H. Raiffa. *Games and Decisions.* John Wiley and Sons, 1957.

[Malone *et ai,* 1988] T. W. Malone, R. E. Fikes, and M. T. Howard. Enterprise: A marketlike task schedule for distributed computing environments. In B. A. Huberman, editor, *The Ecology of Computation.* North Holland, 1988.

[Nash, 1950] J. F. Nash. The bargaining problem. *Econometrica,* 18:155-162, 1950.

[Osborne and Rubinstein, 1990] M. J. Osborne and A. Rubinstein. *Bargaining and Markets.* Academic Press Inc., San Diego, California, 1990.

[Raiffa, 1982] H. RaifTa. *The Art and Science of Negotiation.* Harvard University Press, 1982.

[Rosenschein and Genesereth, 1985] J. Rosenschein and M.R. Genesereth. Deals among rational agents. In *Proc. of the Ninth International Joint Conference on Artificial Intelligence,* pages 91-99, Los Angeles, California, 1985.

[Rubinstein, 1982] A. Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica,* 50(1):97-109, 1982.

[Rubinstein, 1985] A. Rubinstein. A bargaining model with incomplete information about preferences. *Econometrica,* 53(5):1151-1172, 1985.

[Sathi and Fox, 1989] A. Sathi and M.S. Fox. Constraint-directed negotiation of resource reallocations. In L. Gasser and M. N. Huhns, editors, *Distributed Artificial Intelligence, Volume II,* pages 163-194. Pitman/Morgan Kaufmann, London, 1989.

[Selten, 1975] R. Selten. Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory,* 4:25-55, 1975.

[Staid, 1977] I. Stahl. An n-person bargaining game in an extensive form. In R. Henn and 0. Moeschlin, editors, *Mathematical Economics and Game Theory,* Lecture Notes in Economics and Mathematical Systems No. 141. Berlin:Springer-Verlag, 1977.

[Sycara, 1987] K.P. Sycara. *Resolving Adversarial Conflicts: An Approach to Integrating Case-Based and Analytic Methods.* PhD thesis, School of Information and Computer Science, Georgia Institute of Technology, 1987.

[Zlotkin and Rosenschein, 1990] G. Zlotkin and J. Rosenschein. Negotiation and conflict resolution in non-cooperative domains. In *Proceedings of AAAI-90,* pages 100-105, Boston,MA, 1990.