Long Term Constraints in Multiagent Negotiation

Michael Palatnik
Jeffrey S. Rosenschein
Computer Science Department
Hebrew University
Givat Ram, Jerusalem, Israel
mischa@cs.huji.ac.il, jeff@cs.huji.ac.il

Abstract

We consider negotiation over resources among self-motivated agents. Negotiation occurs over time: there are *time constraints* that affect how each agent values the resource. The agents also consider the possibility that they will encounter each other in future negotiations. We cope with agents having incomplete information, emphasizing probability management techniques. Questions arising in this research are: **1.** When is it worth it for an agent to lie? **2.** How beneficial can such lies be? **3.** Is the system with a lying agent robust? **4.** How can lies be discouraged? The main contributions of this work are our ability to deal with multiple encounters among agents, and our treatment of the problem that enables elementary mutual learning.

1 Introduction

Distributed Artificial Intelligence (DAI) is the subfield of Artificial Intelligence (AI) concerned with how automated agents can be designed to interact effectively. Researchers concern themselves both with centrally-designed multi-agent systems (sometimes referred to as the *Distributed Problem Solving* part of DAI), and with multi-agent systems comprising entities that represent diverse interests. We are here interested in this latter model, where agents are *self-motivated* and *rational* (utility maximizing).

The problems analyzed in this paper revolve around the general modeling of task-oriented multi-agent *negotiation*. "Negotiation" has emerged as a basic issue in DAI research. The term has been used in many diverse ways, but generally characterizes communication processes among agents that lead to increased coherent activity.

Our research on negotiation concentrates on the problem of resource allocation and task distribution under time constraints. We also consider the possibility of multiple encounters over time between the same agents. We therefore focus on *long term* considerations in negotiation. Since we wish to cope with the situation where agents have incomplete information, we particularly emphasize techniques for probability management, where each agent has beliefs regarding its current opponent. The questions that arise in this research are: **1.** When is it worth it for an agent to lie? **2.** How beneficial can such lies be? **3.** Is the system with a lying agent robust? **4.** How can lies be discouraged?

The work is an extension of the model of multi-agent negotiations under time constraints presented in [7, 8, 9]. The main contributions of the current work are our ability to deal with multiple encounters among agents, and our treatment of the problem that enables elementary mutual learning.

The problem of resource allocation arises when a set of agents shares a common resource (database, file, printer, network line, etc.) which cannot be used by all interested parties without reaching an agreement. The Task Distribution Problem is the problem of allocating particular tasks (labor portions) to particular agents, such as through market mechanisms. The classical model using this approach is the Contract Net system [2]. Similar problems were examined by Durfee and Lesser [3], by Conry [1], and by Ephrati and Rosenschein [4, 6, 5] (who used voting techniques). Sathi and Fox [12] used a market-like mechanism, where the agents negotiated buying and selling the resources until a compromise is reached.

2 Initial Setting

Our initial assumptions closely follow [9].

1. Bilateral Negotiation — In each given period of time no more than two agents need the same resource. Whenever there is an overlap between the time segments in which two agents need the same resource, these agents will be involved in a negotiation process.

2. Incomplete information — The agents may have no access to their opponents' utility functions, which define "desirability" of the negotiation outcome over time.

3. Rationality — The agents are rational, in the sense that they try to maximize their utilities.

4. Commitments Kept — If agreement is reached, both sides honor it.

5. No Long Term Commitments — Each negotiation stands alone. An agent cannot commit itself to any future activity other than the local schedule reached during the negotiation process.

6. The Threat of Opting Out — The agents can unilaterally opt out when opting out becomes preferable.

7. Common Knowledge — Assumptions (1)–(6) are common knowledge.

We assume that there is a set of agents, that negotiate with each other from time to time on sharing a resource. In distinct negotiation interactions, a given agent may play one of two roles. Either it is attached to the resource while another agent wants to use it—then we say the agent plays the role of A—or it is waiting for the resource used by another agent—then it is playing the role of W. We further assume that there is a finite set of types of agents in the system ($Type = 1, .., k$), each type having a different utility function associated with the negotiation result and the time when it has been reached. The types are associated with some important feature of the agent, whose ultimate meaning is a measure of resource usage. The more frequently an agent uses the system resources, the weaker it is. Later on we will often refer to this feature as the agent's *power*. The types are ordered, in the sense that the first type, 1, is the most heavy user of the system resources. It is denoted by h. Conversely, the lightest user k is denoted by l.

The fact that the agent of type i plays the role of W or A is captured by the notation W_i or A_i respectively. Since information that the agents have is incomplete, we assume that each agent maintains a probability belief with respect to the possible opponent's type. For each agent, ϕ_j, where $j \in Type$, is a probability of its opponent being of type j. Since each agent belongs to one of the types, it is clear that $\sum_{j=1}^{k} \phi_j = 1$ for each agent managing its probability belief. The set of all possible configurations of agents is denoted by $\mathcal{A}$, where $\mathcal{A} = \{W_1, W_2, ..., W_k, A_1, A_2, ..., A_k\}$.

2.1 Model Structure

Here we briefly review the basics of the Alternative Offers model. For the detailed analysis see [10]. Two agents will negotiate in order to divide M ($M \in \mathbb{N}$) units of a common resource.

We assume that the negotiation moves are performed at the discrete time moments $t \in \mathcal{T}$, where $\mathcal{T} = \{0, 1, 2, ...\}$. At a given moment $t \in \mathcal{T}$ an agent makes an offer $s \in S$. At the same time period, its partner responds either by accepting the deal (**Yes**) or rejecting it (**No**), or opting out (**Opt**). If the offer was rejected, the opponent has to make its counter offer in the next time period $t+1$. If the offer was accepted, it must then be fulfilled. The third outcome means that the encounter is over.

Practically, when dealing with the resource allocation problem, we assume that unilateral opting out may lead to the unconventional usage of the resource, which in turn may cause the resource to temporarily shut down.

Generally, by a **strategy** of a player in an extensive game[1] the game theorists mean a function or automata that specifies an action at every node of the tree at which it is the player's turn to move.

The outcome (s, t) of a negotiation is the reaching of agreement s in period t; the outcome (**Opt**,t) means that one of the agents opts out of the negotiation at the time period t; and the symbol **Disagreement** indicates a perpetual disagreement. We also assume that agent $i \in \mathcal{A}$ has a continuous utility function over all possible outcomes: $U^i\colon \{\{S \cup \{\mathbf{Opt}\}\} \times \mathcal{T}\} \cup \{\mathbf{Disagreement}\} \rightarrow \mathbb{R}$. The main factor that plays a role in reaching an agreement in a given period t is the worst agreement $\hat{s}^{W,t} \in S$ which is still preferable to W over opting out at t.

We want to find a strategy that will guarantee an agent the optimal outcome in terms of utility. When the information is incomplete, we use the notion of *sequential equilibrium*,

[1]The *extensive form* of a game specifies when each agent will make each of the choices facing him, and what information will be at its disposal at this time.

arguing that if a set of strategies is in sequential equilibrium, no agent will choose a strategy from out of the set.

In our model, **Sequential Equilibrium** is a sequence of $2k$ strategies (one for each agent $A_1, A_2, .., W_1, .., W_k$) and a system of belief, so that each agent has a belief about its opponent's type. At each negotiation step t the strategy for agent i is optimal given its current belief (at step t) and its opponent's possible strategies. At each negotiation step t, each agent's belief (about its opponent's type) is consistent with the history of the negotiation. We assume each agent in an interaction has an initial probability belief.

Initially we impose three conditions on the sequence of strategies and the agent's system of beliefs.

1. Sequential Rationality — The optimality of agent i's strategy after any history h depends on the strategies of $W_1, .., W_k$ and on its beliefs $p_i(h)$. That is, an agent i tries to maximize its expected utility with regard to the strategies of its opponents and its beliefs about the probabilities of its opponent's type after the given history. It does not take into consideration possible interactions in the future (this last requirement will be relaxed in Section 3).

2. Consistency — Agent i's belief $p_i(h)$ should be consistent with its initial belief $p_i(\emptyset)$ and with the possible strategies of its opponent. An agent must, whenever possible, use Bayes' rule to update its beliefs.

3. Never Dissuaded Once Convinced — Once an agent is convinced of the type of its opponent with probability 1, or convinced that its opponent can't be of a specific type, i.e., the probability of this type is 0, it is never dissuaded from its view.

2.2 Assumptions

This section summarizes all assumptions mentioned above, as well as other assumptions we make within the initial setting.

1. Agent A prefers disagreement over all other possible outcomes, while agent W prefers any possible outcome over disagreement.

2. For agreements that are reached within the same time period, each agent prefers to get a larger portion of the resource.

3. For any $t_1, t_2 \in \mathcal{T}; s \in S$ and $i \in \mathcal{A}$ if $t_1 < t_2$, $U^W((s, t_1)) \geq U^W((s, t_2))$ and $U^A((s, t_1)) \leq U^A((s, t_2))$

4. Agent W prefers to obtain any given portion of the resource sooner rather than later, while agent A prefers to obtain any given portion of the resource later rather than sooner. However the exact values c_W and c_A are private information. Each agent knows its own value c, but it may not know its opponent's c, though it knows that it is one of k values— depending on the opponent's type t. W_l loses less than W_h, while waiting for the resource, and at the same time A_l gains less than A_h, while using the resource. We assume that this is common knowledge.

5. W prefers opting out sooner rather than later; A prefers opting out later rather than sooner (since A gains over time while W loses over time).

6. If there are some agreements that agent W prefers over opting out, then agent A also prefers at least one of those agreements over W's opting out even in the next period.
7. For all $i, j \in \mathcal{A} \quad U^i((\hat{s}^{j,0}, 0)) \geq U^i((\hat{s}^{i,0}, 0))$.

The last assumption ensures that an agreement is possible at least in the first period; there is an agreement that both agents prefer over opting out. Having defined the domain and the model, which closely follow Kraus et al.'s setting, we proceed with our own further exploration of the model.

3 One Lying Agent

Let's consider what happens if a liar is introduced among the agents behaving according to the above model. A "liar" is an agent designed to benefit over time, but not necessarily immediately.

We adopt the assumptions 1–7. Notice that all three assumptions 3, 5, and 7 do relate only to the local passage of time. That is, time does cost something *within* an interaction: the "stop-watch" is started when an agent wants to receive the resource (receives a new task), and is stopped when the resource is no longer needed (the task is done). However, when speaking about "long term" considerations, we may speak about utility "inflation." We will introduce a new assumption that captures this notion. First we introduce global time, advancing in discrete time moments $\tau \in T$, where $T = \{0, 1, 2, \ldots\}$.
Assumption 8. Utility value over time
Utility value decreases over time with a constant discount rate δ, $\delta \in [0, 1]$.

$$U(\tau_0 + \Delta\tau) = \delta^{\Delta\tau} U(\tau_0).$$

Here $U(\tau_0)$[2] is the utility of an event which occurs at the current moment. $U(\tau_0 + \Delta\tau)$ is the utility of the same event that would occur after a time delay $\Delta\tau$.
For any $\tau \in T$, $i \in \mathcal{A}$ and $s \in S$: $U^i(s) \times (\delta^{\tau} - \delta^{\tau+1}) \ll |c_i|$.

Def 1 The Strongest the Agent believes that the Opponent thinks him to be
Let's say the agent W_i can estimate what is the strongest type that A_j still believes W_i to be. Thus, it is the maximal $n \in$ Type such that $\phi_n^{A_j} \neq 0$, according to W's estimation. We'll call this number γ.

Let's examine what will be beneficial lies. Considering the "single-minded liar" (or liar for the sake of lies), we can intuitively see that in the extreme case where the number of types in the system is very large (tends to infinity), the strongest type is never reached. Consequently the "single-minded liar" gets stronger and stronger in A's belief, but it never enjoys its power.

[2] We'll denote it also as U_0. Note that calculating its future utility, each agent may actually start with the current moment as if it were $\tau = 0$.

The rationally designed W_i shouldn't try to convince its opponent that it is the strongest type in the system, rather it should maximize its utility function. It should get enough interactions to enjoy its lies.

Assumption 9. Estimation of Number of Interactions
Agent W knows the lower bound $Inf(I)$ on the number of its interactions with the specific agent A_j during the life of the system.[3]

Assumption 10. Estimation of the Interaction Time
Agent W knows the upper bound $Sup(\Delta\tau)$ of the time interval between two successive interactions.

Let's investigate the difference to agent W's Utility in the case where it tells the truth (i.e., as proposed in [9]) and in the case where it lies. We'd like to know two conditions: [1] the utility difference is positive (weak), [2] the utility is maximal (strong).

Let's consider the simplified case where there are no other interactions in the system. That is, the only encounters are between W_i and A_j. Following the results of [9] — Theorem 3 — we consider each interaction to consist of four phases. First, W makes its "ritual" offer — A rejects — A in turn offers a deal — W either opts out or accepts it. Suppose in the first encounter A_j offers the bound deal $\hat{s}^{W_{n'},1}$.[4] Here, $\gamma = n'$. Suppose also W wants to convince A that it is as strong as m, $m > n'$. The maximal number of opts needed to satisfy W's goal (if it is at all possible) is $Sup(\#opt) = m - n'$, that is at most $m - n'$ interactions, i.e., the minimal (worst) utility W will have whenever its goal is satisfied, $U_1^{W_{lies}}$, is given by the following equation: $U_1^{W_{lies}} = \sum_{i=0}^{m-n'} U^W((Opt, \Delta\tau_i)) = U_0^W(Opt) \sum_{i=0}^{m-n'} \delta^{\Delta\tau_i}$, where $\Delta\tau_i$ is the delay of the i^{th} encounter starting from the current time. Thus

$$U_1^{W_{lies}} = U_0^W(Opt) \sum_{i=0}^{m-n'} \delta^{\Delta\tau_i} \tag{1}$$

Similarly, after $m - n'$ encounters are over, the lying agent earns

$$U_2^{W_{lies}} = U_0^W(\hat{s}_W^m) \sum_{i=m-n'+1}^{I+1} \delta^{\Delta\tau_i} \tag{2}$$

The total utility of W_i pretending to be m, during the life of the system, is the sum of the two components: $U^{W_{lies}} = U_1^{W_{lies}} + U_2^{W_{lies}}$.

Figure 1 depicts behavior of either a truth-telling or a lying agent W_i in different settings. The axes represent the discrete set $Type$. For both truth-telling and lying agents there exist two distinct situations.

For the truth-teller — (1) $i > \gamma$ and (2) $i \leq \gamma$.
For the liar — (1) $m > \gamma$ and (2) $m \leq \gamma$.
The liar's utility described above belongs to the cases a and b of the figure. The case c with respect to the liar is examined below.

[3] For simplicity, we still will refer to I even if what is actually known is its lower bound.

[4] Since we are frequently referring to the deal offered by agent A in a specific interaction, with the successive W's response, throughout the rest of the work we will omit mentioning $t = 1$. Thus $U(Opt) \equiv U((Opt, 1))$ and $U(\hat{s}_W^i) \equiv U((\hat{s}_W^{i,1}, 1))$.

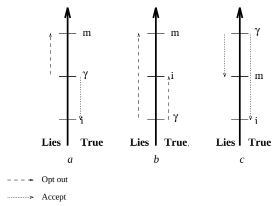

Figure 1: Truth Telling and Lying Behavior in Different Settings

The maximal utility of truth-telling agent W_i (cases a and c) is composed of three components:
1. $U_1^{W_{truth}}$ — W accepts the offers of A in a decreasing sequence.
2. $U_2^{W_{truth}}$ — W opts out, when the offer of A doesn't satisfy it.
3. $U_3^{W_{truth}}$ — W accepts all the offers, which exactly fit its power, until all the interactions are over.

$$U^{W_{truth}} = U_1^{W_{truth}} + U_2^{W_{truth}} + U_3^{W_{truth}}$$

$U_1^{W_{truth}} = \sum_{k=0}^{n'-i} U_0^W(\hat{s}_W^k)\delta^{\Delta\tau_k}$; $U_2^{W_{truth}} = U_0^W(Opt)\delta^{\Delta\tau_{n'-i+1}}$; and $U_3^{W_{truth}} = \sum_{k=n'-i+2}^{I} U_0^W(\hat{s}_W^i)\delta^{\Delta\tau_k}$. i here is W_i's actual power.

Thus, after the first offer, agent W may estimate its benefit in terms of
1. Total number of interactions W_i — A_j, I.
2. The time delay $\Delta\tau_k$ of the kth interaction.
3. Supposed *ambition* (the power that W wishes to convince A it is).
Similar results can be obtained if ambition $m \leq$ initial offer n'.

In general, comparing the two utilities — max $U^{W_{lies}}$ and $U^{W_{truth}}$ — W can always decide what is more beneficial. However a bit more elegant way would be to deal with a single equation, pointing out the type that will guarantee (within the estimation) the best behavior. The maximization of the unified function will output the true type of W i whenever the best alternative is to behave according to it. The equation distinguishes between the two cases: when the ambition $m > \gamma = n'$ and when $m \leq n'$. In the latter case (case c in the Figure 1) $U^{W_{lies}}$, similarly to $U^{W_{truth}}$, is decomposed into three components:

$$U^{W_{lies}} = \sum_{k=0}^{n'-m} U_0^W(\hat{s}_W^k)\delta^{\Delta\tau_k} + U_0^W(Opt)\delta^{\Delta\tau_{n'-m+1}} + \sum_{k=n'-m+2}^{I} U_0^W(\hat{s}_W^i)\delta^{\Delta\tau_k}$$

It is clear that to satisfy the strong requirement W has to maximize the utility difference ΔU, which cannot be done unless it has an estimation of the total interaction number I. Knowledge of $Sup(\Delta\tau)$ gives an estimation of each $\Delta\tau_k$. Since it provides the upper bound, W is guaranteed that it has a positive profit, though this profit is optimal only within a certain precision.

Following are some interesting special results (confirmed through computer simulation). Assuming $U_0^W(Opt)$ is small ($U_0^W(Opt) \ll U_0^W(\hat{s}_W^m)$), utility does not decrease with the global time ($\delta = 1$) and $U_0^W(\hat{s}_W^m)$ is proportional to the type $m \in Type$, the best ambition m for a type $i < I/2$ is given by $m = min(I/2, |Type|)$. Subsequently, if the number of interactions is considerably large ($I/2 > |Type|$), each agent W_i, $i < I/2$ will behave as the strongest agent in the system, while each agent W_i, $i \geq I/2$ will tell the truth. If discount rate δ has a reasonable value, all the weak agents will pretend to be the same power m. Beginning from this m ($i \geq m$) all the agents will tell the truth.

Suppose we decrease the value of δ gradually. Beginning at some value the truth becames preferable for everybody. Actually, it is a special case when "inflation" becames so large that no long term considerations make sense. From another perspective, it is the special case of m described in the previous paragraph. With the growth of the discount rate (fall of δ), m tends to 1.

4 How to Make the Lies Unprofitable

A good way to return the strategies to sequential equilibrium would be to ensure that it is unprofitable to lie. For this to hold, A's system of belief should have a way to know that one of a number of opponents does lie. Then it could make a backward step assigning to all the candidates at the same place (power, type) less credit. To allow agent A to calculate that there is a liar we are obliged to introduce some interdependence between the opponents in agent A's system of belief. So we'll try to develop the second way of managing the system of belief briefly outlined in the previous section.

4.1 Matrix of Belief

Suppose we have $N + 1$ agents in the system. Let's consider an agent A_j, who thus has N potential opponents. Suppose also there are k types of agents in the system ($|Type| = k$). Before any interactions have started, A_j has an initial probability distribution, which has a simple intuitive meaning: A knows the precise number of agents of each type in the system. Let's introduce the *matrix of belief*, which is very similar to the system of beliefs in the initial setting. For simplicity's sake, we consider that A can enumerate its opponents, so that it can identify them during successive encounters.

Assumption 11. A Limit per Type
Agent A knows what is the number of its opponents for each possible type in the system. We'll denote this number N_j, where $j \in Type$.

Def 2 Matrix of Belief *The matrix of belief of agent A is a $N \times k$ matrix, where N is the number of A's opponents and k is the number of types in the system, consisting of probability assignments $\phi_{i,j}$. Each probability $\phi_{i,j}$ has a clear meaning of probability of an opponent i to be of type j.*

Def 3 Consistency of the System of Belief *We'll say that the matrix of belief is maintained in a consistent way if at every period of time $t \in \mathcal{T}$ the following constraints hold: $\sum_i \phi_{i,j} = N_j$, and $\sum_j \phi_{i,j} = 1$.*

		Types		
		0	1	2
Opponents	I	0.5	0.5	0
	II	0	0	1
	III	0.5	0.5	0

Figure 2: An example of a matrix of belief

The matrix of belief is revised each time A interacts with some agent W_i. As was proved by Kraus et al. (Theorem 3 in [9]), each interaction follows the same course. Agent W makes an offer, A rejects and makes its counter offer, then at the final stage W either accepts A's offer or opts out. The offer made by A has a clear meaning of the last acceptable offer of an opponent W of a specific type g (*guess*). W's precise behavior depends on whether g is weaker than its actual type, or it is stronger or equal. Now, both opting out and acceptance of A's offer gives A information about W's actual type.

From the informational point of view, we can distinguish between a *productive* encounter and a *fruitless* one. This distinction indicates whether the interaction has given new information to A's belief matrix. In the case when the interaction was productive, A will modify its matrix in a consistent way.

If such a modification which preserves consistency can't be managed or, in other words, the consistency constraints are violated, we say that there is a **conflict**. A conflict's true meaning is that there is a *liar* in the system. Once the liar's presence is discovered, A will try to minimize its impact, using a sort of "backtracking" mechanism.

We'll say that the matrix of belief *converges* if after a finite time period $t \in \mathcal{T}$ no interaction with any opponent W yields a modification of the matrix. We'll say that the matrix *absolutely converges* if the matrix (1) converges and (2) contains no probability assessments except 0 and 1. Finally we'll say that the matrix of belief does not converge if there is no finite time period $t \in \mathcal{T}$ such that interaction with any opponent W will not have any impact on the probability assessments.

Now let's see more exactly how A can maintain the matrix of belief in a proper way. Let's consider the simplest variant of a matrix of belief, where the number of types in the system is equal to the number of A's opponents: $N = |Type|$. We claim also that there is no reason to introduce into the matrix the columns corresponding to the types which, according to A's initial knowledge, do not appear in the system. Thus, in our simplest case, A's initial matrix of belief is 3×3 matrix, where the sum of each column as well as the sum of each row is equal to 1. The final state of the matrix, which we are eager to reach if the matrix does converge, is the matrix consisting only of zeros and ones. We'll consider each possible final state to be a distinct hypothesis. Clearly, the number of hypotheses in a square $N \times N$ matrix is $N!$. Now we will represent each hypothesis as an $N-$tuple, where N is the number

#	*tuple*		
1	1	2	3
2	1	3	2
3	2	1	3
4	2	3	1
5	3	1	2
6	3	1	2

Figure 3: Initial Hypothesis Set for 3×3 matrix

0.33	0.33	0.33
0.33	0.33	0.33
0.33	0.33	0.33

Figure 4: Initial probability assignment for 3×3 matrix

of the opponents and the k^{th} number corresponds to the type of k^{th} A's opponent. So the range of each member of the tuple is $[1 \ldots k]$. See Figure 3 for the full set of hypotheses for our case.

Proceeding from the *Principle of Indifference* we will assign each hypothesis from the hypothesis set H an equal probability. Now, since the hypothesis set H forms the full set of hypotheses (they are disjoint and cover all the possibilities), we will assign to each of them the equal initial probability: $p(H_i) = \frac{1}{N!}$.

The next question to answer is how does A, having a probability assignment for each possible hypothesis, compute the probabilities of each elementary event (i.e., the values in the squares). The probability of each elementary event (that is, opponent i being of type j) is a relative quantity of this particular event in the total amount of hypotheses. Each square (i, j) will receive a probability $\frac{m}{|H|}$, where $|H|$ is the total number of hypotheses in the set and m is the number of this square's occurrences within the hypothesis set. For instance, in our example the square $(1, 1)$ is mentioned exactly 2 times — these are the 1^{st} numbers in the first two hypotheses H_1 and H_2. Since the total number of hypotheses is $3! = 6$, the square $(1, 1)$ will receive the probability assignment $\phi_{1,1} = \frac{2}{6} = \frac{1}{3}$. Figure 4 depicts the initial probability assessment in the case being considered. The intuition underlying this probability assignment procedure is clear. First, all hypotheses, as mentioned, are equally probable. Second, a realization of a certain hypothesis means the squares it passes through are also being realized with the same probability.

Each productive interaction inserts a zero in one or more squares in the matrix of belief. We will call these zeros *primary zeros* (as opposed to zeros received as a result of a calculation). Moreover, we want to distinguish between *left* and *right* primary zeros. The left ones are the zeros which mean the opponent can't be as *weak*: it has opted out at some stage, showing that it is not weak like this. The right zeros mean the opponent is not as strong. The zeros of this type appear after an agent has accepted an offer, which means the deal does satisfy it. The left zeros vs. the right ones can't appear as a result of a "pretending" (lying) action, since it doesn't make sense for an agent to receive a worse deal than it could.

A zero in a square clearly means that a hypothesis can't pass through it. Thus, after a productive interaction, A removes all the hypotheses that became impossible and then recomputes the probabilities. Assuming there is no liar in the system, the desired convergence state (the state of zeros and ones) is reached when there remains only one hypothesis in the set. On the other hand, the described procedure provides us with a clear criteria of *conflict*. Conflict is reached when there are no more hypotheses in the set.

The matrix of belief, maintained as described, is consistent in the above defined sense.

Lemma 1 The Matrix of Belief's Consistency
If the matrix of belief is maintained as described above it is consistent unless a conflict occurs.

Proof: The proof can be found in [11].

We have already mentioned that in the case of a square matrix ($N = k$; $\forall j N_j = 1$) the initial number of hypotheses is $|H_{init}| = N!$. In the general case this number is defined by the following equation.

$$|H_{init}| = \prod_{j=0}^{k} \binom{N_j}{N - \sum_{l=0}^{j} N_{l-1}} \tag{3}$$

4.2 Backtracking

The state of conflict is solved by means of one step of **backtracking**. The intuition underlying the backtracking procedure is the following. If the state of belief of an agent A contradicts its domain (system) knowledge, at least one of the opponents is lying, claiming that it is more strong than it really is. Practically, it means that one (or more) of the left primary zeros is a wrong one and as such should be disregarded. But A has no idea who among its opponents is the liar, therefore it disregards all the rightmost left primary zeros. This is accomplished by returning to the hypothesis set all these hypotheses, which now become valid. Completing one step of the backtracking and recomputing all the elementary event probabilities, A gets the consistent matrix of belief.

Note that one productive interaction may imply more than one primary zero. Introducing them one by one we may easily catch the conflict state, and eliminate its cause (one step of the backtracking). Then A will proceed, introducing the remaining primary zeros.

Clearly, when using backtracking, if the matrix of belief of agent A did absolutely converge without a liar, it will not converge in a system with a liar. Thus the liar has no way to benefit from its lies, even though it can prevent the other agents—namely, those who have the same power that the liar pretends to be—from receiving optimal utility.

Note that in adopting the matrix of belief and the hypotheses-based procedure, we save all the history by means of the hypothesis set. After the backtracking procedure, we open a new "sub-history" which is consistent with all of the previous history, except the "candidate-for-liars" opting outs.

5 The Optimal Offer

Going back to Theorem 3 of [9], we notice that A makes the choice of the deal, offered to the given opponent by maximization of its expected utility. It makes a lot of sense considering the local short-term view, but appears to be not always justified with regard to multiple encounters.

It occurs that offering the deal maximizing the expected utility, A often gets in a sort of trap. We use the term *trap* to denote situations where A's matrix of belief does converge but doesn't converge absolutely. It is also remarkable that these situations do not depend on the presence of the liar in the system, but are rather influenced by A's utility function configuration. We will illustrate our observations with an example.

Example 1 Consider the case where A has a 4×4 matrix of belief. The opponent 1 is of type 1, 2, etc. A has the following utility function:

Outcome	*Utility*
1	5
2	4
3	3
4	2
Opt	1

Table 1: Utility Function

Here the left column describes the possible outcomes—the deal offered as if the opponent is of type 1, 2, etc. and opting out, while the second column provides the utility values. Consider the following state of the matrix of belief:

0.25	0.25	0.50	0.00
0.50	0.50	R0	R0
0.25	0.25	0.50	R0
L0	L0	L0	1.00

Figure 5: Belief Matrix

Here 'L0' denotes the left primary zero, whereas 'R0' denotes the right primary zero. Now we refer to this state as a trap, since no interaction will carry new information—i.e., a primary zero. Thus A here is stuck: the matrix has converged but not in an optimal way.

A straightforward proposal would be for A to just avoid the traps. That means whenever the maximum of the expected utility falls onto the rightmost (the strongest among the believable) type, offer the next to it (last but one). Somewhat surprisingly, this approach works. However we will try to exploit another approach to the problem.

Let's consider the information contributed by both possible responses (Acc, Opt) of the opponent W. We will represent this information in terms of utility. We will denote the maximal believable (the strongest) type n_{max}, the minimal — n_{min}, and the type chosen by A to make the offer n and n_{off}. Consider an offer made from the middle of the type scale. Figure 6 depicts the situation when the response is either *Accept* or *Opting Out* from the informational standpoint. In both cases the gray area shows the segment of the types which becomes improbable. In case of *Accept* each offer from within the gray area will provide A with utility less than $U(n)$. Thus we may say (informationally) that A spares a number of discounted refreshes equal to the number of the types within the gray area.

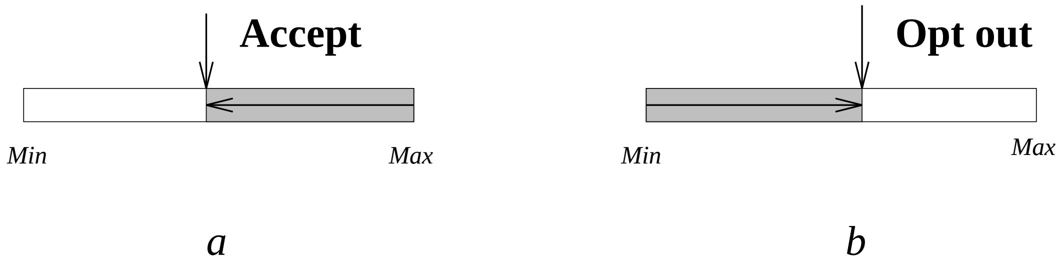

Figure 6: A Making an Offer

The next equation gives us the total utility A enjoys whenever W accepts its offer $\hat{s}^{W_n}$.[5]

$$U_{Acc}(n) = \sum_{i=0}^{n_{max}-n-1} [U(n) - U(n_{max} - i)]\delta^{\Delta\tau_i} \;+\; U(n) \tag{4}$$

$\Delta\tau_i$ is the delay of the i^{th} interaction from the moment of the current decision.

Speaking of the case when W opts out we notice that for each type from within the gray area (Figure 6, b) A spared at least[6] $U(n+1) - U(Opt)$, where the total number of the spared opts is equal to $1 + n - n_{min}$. When W opts out, A receives the following utility.

$$U_{Opt}(n) = \sum_{i=0}^{1+n-n_{min}} [U(n+1) - U(Opt)]\delta^{\Delta\tau_i} \;+\; U(Opt) \tag{5}$$

Finally we propose that A, instead of maximizing the plain expected utility, should maximize the expected utility in sum with expected *information*. The probability of offer $\hat{s}^{W_n}$ being accepted is $P_{Acc}(n) = \phi_{n_{min}} + \phi_{n_{min}+1} + \ldots + \phi_n$. The probability W opts out in response to $P_{Opt}(n) = 1 - P_{Acc}(n)$.

$$n_{off} = \arg\max_n [U_{Acc}(n)P_{Acc}(n) + U_{Opt}(n)P_{Opt}(n)] \tag{6}$$

[5]Throughout this section, for simplicity's sake, instead of $U_0^A((\hat{s}^{W_n,2}, t))$ we refer to $U(n)$.
[6]It should be clear that this profit may not be immediate.

Example 2 Let's examine the previous example using the new rule for generating the offer. For simplicity we'll take $\delta = 1$ (no discount). The 1^{st} and the 3^{rd} row of the matrix of belief (Figure 5).
For $n = 1$ $F = 1/4[(5-3)+(5-4)+5] \ + \ 3/4[(4-1)+1] = 2+3 = 5$.
For $n = 2$ $F = 1/2[(4-3)+4] \ + \ 1/2[(3-1)+(3-1)+1] = 2.5+2.5 = 5$.
For $n = 3$ $F = 1 \times 3 + 0 = 3$.
Thus n_{off} =1 or 2. The choice in the case when the maximum is not unique is up to the implementor. For the second row:
$n = 1$: $F = 1/2[(5-4)+5] + 1/2[(4-1)+1] = 5$,
$n = 2$: $F = 1 \times 4 = 4$.
In both cases A doesn't fall into the trap.

Type	1	2	3	4
ϕ_{Type}	1/4	1/4	1/2	0
P_{Acc}	1/4	1/2	1	
P_{Opt}	3/4	1/2	0	Opt
Utility	5	4	3	1

Figure 7: The Data used in the Example

Let's examine if computing the optimal offer in the proposed way A is guaranteed not to fall into traps. It is intuitively clear that the rightmost type provides A with zero information. Since the importance of the information vs. current profit grows with the increase of the long-term considerations, the less A is concerned about its future the more probable it is that it can get into a trap. We'll show formally when A will prefer to offer the strongest type offer over the "last but one." Let $P_{Acc}(n_{max}-1) = p$. For the above statement to hold, the following inequality should be true:

$$U_{Acc}(n_{max}) - U_{Acc}(n_{max}-1)p > U_{Opt}(n_{max}-1)(1-p) \iff$$

$$U_{Acc}(n_{max}) - U_{Opt}(n_{max}-1) > (U_{Acc}(n_{max}-1) - U_{Opt}(n_{max}-1))p \qquad \text{(a)}$$

Since $p < 1$ (and $U_{Acc}(n_{max}-1) > U_{Acc}(n_{max})$), this holds only when $U_{Acc}(n_{max}) > U_{Opt}(n_{max}-1)$ (b). $U_{Acc}(n_{max}) = U_{min}$, that is the minimal utility A can receive in the given setting. (b) $\iff$ $U_{min} > \sum_i^{1+n_{off}-n_{min}}[U_{min} - U(Opt)]\delta^{\Delta\tau_i} + U(Opt)$. If we assume that $U(Opt) \ll U_{min}$, this last equation is likely to hold when δ tends to 1, that is, the importance of the current profit becames crucial. However, even when (b) holds, condition (a) also has to be true. The greater the value of p, the less likely it is that (a) will hold. Notice also that $p = \phi_{n_{min}} + \phi_{n_{min}+1} + \ldots + \phi_{n_{max}-1}$, which means that $p \to 0$ is a very special case.

6 Conclusion

This work develops the strategic model of negotiation in the environment of multiple automated agents sharing either common resources or common tasks. We investigated the protocol based on the model of Alternative Offers. The innovative element of the research is the introduction of long-term considerations involving both global (outside a single encounter) passage of time and multiple encounters. Considering only the case of incompletely informed agents, we presented a technique to maintain the consistency of beliefs. The latter was based on a generalization of Bayes' rule to the case of a two-dimensional matrix. Future work will include considering negotiation among more than two agents, and situations in which the power of agents changes over time.

References

[1] S. E. Conry, R. A. Meyer, and V. R. Lesser. Multistage negotiation in distributed planning. In Alan H. Bond and Les Gasser, editors, *Readings in Distributed Artificial Intelligence*, pages 367–384. Morgan Kaufmann, San Mateo, California, 1988.

[2] R. Davis and R. G. Smith. Negotiation as a metaphor for distributed problem solving. *Artificial Intelligence*, 20(1):63–109, 1983.

[3] E. H. Durfee and V. R. Lesser. Using partial global plans to coordinate distributed problem solvers. *Proceedings of IJCAI-87*, pages 875–883, 1987.

[4] E. Ephrati and J. S. Rosenschein. Constrained intelligent action: Planning under the influence of a master agent. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, pages 263–268, San Jose, California, July 1992.

[5] E. Ephrati and J. S. Rosenschein. Distributed consensus mechanisms for self-interested heterogeneous agents. In *First International Conference on Intelligent and Cooperative Information Systems*, pages 71–79, Rotterdam, May 1993.

[6] E. Ephrati and J. S. Rosenschein. multi-agent planning as a dynamic search for social consensus. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pages 423–429, Chambery, France, August 1993.

[7] S. Kraus and J. Wilkenfeld. The function of time in cooperative negotiations. In *Proc. of AAAI-91*, pages 179–184, California, 1991.

[8] S. Kraus and J. Wilkenfeld. Negotiation over time in a multi-agent environment: Preliminary report. In *Proc. of IJCAI-91*, pages 56–61, Australia, 1991.

[9] S. Kraus, J. Wilkenfeld, and G. Zlotkin. Multiagent negotiation under time constraints. Technical report, University of Maryland, College Park, MD 20742, 1992.

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

[11] Michael Palatnik. Long term constraints in multiagent negotiation. Master's thesis, Hebrew University, 1993.

[12] A. Sathi and M. Fox. Constraint-directed negotiation of resource reallocations. In L. Gasser and M. Huhns, editors, *Distributed Artificial Intelligence*, volume 2, chapter 8, pages 163–193. Pitman, London, 1989.