A Strategic Negotiations Model with Applications to an International Crisis

Sarit Kraus*
Department of Mathematics and Computer Science
Bar Ilan University
Ramat Gan Israel 52900
e-mail: sarit@bimacs.cs.biu.ac.il

Jonathan Wilkenfeld
Dept. of Government and Politics
University of Maryland
College Park, MD 20742
(301)-405-4160 e-mail:jwilkenf@umdd.umd.edu

Note: This paper appears in *IEEE Transactions on Systems, Man, and Cybernetics*, 23(1):313-323 1993.

Abstract

The area of automated negotiation has been of particular interest in AI due to the important role negotiations play in facilitating understanding and the achievement of cooperation among entities with differing interests, whether they be individuals, organizations, governments, or automated agents. This paper presents a strategic model for negotiation of alternative offers, with specific application to international crises. In this model, both players can opt out, and while one loses over time, the other gains (up to a point). Specific issues are: conflicting objectives and utility functions of parties and the impact of time on bargaining behavior in crisis. The general model has relevance to the hostage crisis from which it was built, and subsequent applicability in building an automated negotiation agent for experimental and training purposes.

*This material is based upon work supported in part by the National Science Foundation under Grant No. IRI-9123460. Part of this work was completed while the author was visiting the Institute for Advanced Computer Studies, University of Maryland, College Park and the Department of Computer Science, Hebrew University, Jerusalem

1 Introduction

The negotiation process facilitates understanding and the achievement of cooperation among entities with differing interests, whether they be individuals, organizations, governments, or automated agents. Our long term objective is the creation of a prototype automated negotiator, as part of the development of a simulation environment of a real world situation, in which negotiators can be trained and where experiments can be conducted. As a basis for the development of such a simulation environment, we have developed a formal theory of negotiations in order to analyze the negotiation situation and to determine the best strategies.

This paper presents a strategic model of negotiation, with specific application to international crises as the first step in developing such a theorem. Our work is based primarily on Artificial Intelligence concepts. Related work in bargaining and negotiation theory, in the general realm of economics and game theory, and foreign policy analysis and crisis decision-making within the domain of political science, are suitably modified for use in an Artificial Intelligence approach.

The specific issues in the model are the conflicting objectives and utilities of the parties and the impact of time on bargaining behavior. While the theoretical discussion and axioms and proofs apply to a general case of negotiation, the theory was built by focusing on a specific hypothetical negotiation between real world international actors. In the process of formalizing the behavior of specific actors, in consultation with regional and negotiation specialists, we identified areas for generalization; the resultant general model has relevance to the specific case upon which it was based, as well as subsequent applicability in building an automated agent as a participant in a simulation of this case, and beyond this specific case to the general class of crisis negotiations.

We begin by examining previous work in the fields of distributed artificial intelligence, negotiation agents, and game theory. A brief description of the hostage crisis follows, serving as the substantive grounding for the theoretical work. We introduce the strategic negotiation model and review its central definitions, theorems, and proofs. Finally, we revisit the hostage crisis in the light of the strategic negotiation model we have developed.

2 Previous and Related Work

In this section we will briefly review some of the related work in the areas of Artificial Intelligence, bargaining and negotiation theory, and crisis analysis.

2.1 Previous Work in Artificial Intelligence

The study of multi-agent interaction has been receiving increasing attention within Artificial intelligence (AI). This is a direct outgrowth of the serious consideration currently being given to agents operating in challenging, real-world environments. For many years, highly restricted domains were considered sufficient for AI research purposes, and agents such as Shakey [12] could be designed and built for operation in simplified, restricted environments.

The research on agent architectures and on planning typically made several standard assumptions, including the existence of a static domain, the lack of deadlines, and the existence of a single agent, i.e., *our* agent. Once researchers began, for a variety of reasons, to move into realistic domains, these assumptions had to be quickly discarded. The research in planning and agent architectures of the last decade has been focused precisely on the transformation of single-agent, atemporal, static theories into multi-agent, temporal, dynamically capable ones.

A community of researchers working on distributed artificial intelligence (DAI) has arisen (for a survey of DAI see [1, 15]). One of the most difficult subjects that has occupied the efforts of the DAI community has been the subject of negotiation [55, 19, 59, 41, 5, 9, 47, 51, 34, 64, 33, 11, 35, 61].

Davis and Smith's work on the Contract Net [55] introduced a form of simple negotiation among cooperative agents, with one agent announcing the availability of tasks and awarding them to other bidding agents. Malone refined this technique considerably by overlaying it with a more sophisticated economic model [41], proving optimality under certain conditions. While Davis and Smith's original work assumed some autonomy among agents, these agents willingly bid for tasks without explicit motivation. Malone's work introduced a motivational framework in the language of economic theory, and at the same time provided a more theoretical language in which to discuss the task-sharing algorithm.

These efforts in DAI and others that have followed dealt with negotiations in the case of cooperative systems which are designed to achieve a common general task, or in which the agents belong to the same organization or unit (see for example [19] which describes a method for synthesizing multi-agent plans from simple single-agent plans, [52] which deals with project management, [9] and [10] which deal with the vehicle monitoring domain, and [51] which deals with resource reallocations). Conflicts among the agents in these environments may arise while each tries to achieve its own sub-tasks (for example, they may need to share the same resources), but their overall task is the same.

Our work takes as a point of departure the work of researchers who have studied the negotiations that could take place among agents that serve the interests of truly distinct parties [59, 47, 63, 25, 30, 31]. The agents are autonomous; they have their own utility functions, and no global notion of utility plays a role in their design. The agents are *individually motivated.*

For example, Sycara [59] presented a model of negotiation that combines case-based reasoning and optimization of the multi-attribute utilities of the agents. She implemented her ideas in a computer program called the PERSUADER which resolved adversarial conflicts in the domain of labor relations, and tested her system using simulations of such domains. While she concentrated on the perspective of the mediator (see also [23]), we want to analyze such situations from the point of view of the autonomous agents that participate in the conflict, and to concentrate on the time constraints of the situations.

Rosenschein and Genesereth [47] used certain game-theoretic techniques to model communication and promises in multi-agent interaction. There, the process of negotiation was severely restricted; the agents could only make single, simultaneous offers. This work was extended by Zlotkin and Rosenschein in [63]. Using game theoretic results (mainly of Harsanyi [21]), they introduced a negotiation protocol for the case of agents who are able to share a discrete set of tasks with one another. In their model the impact of the passage of time in the negotiation is not taken into consideration, and they assume that in each step at least one of the agents has to make a concession, otherwise conflict results.

Other extensions of this models were published in [64]. Comparing this work to ours, we make almost no assumptions about the protocol the agents use for negotiations. Also, our model takes the passage of time during the negotiation process itself into consideration, which in turn influences the outcome of the negotiations and avoids delays in reaching an agreement.

Matwin et al. [42] developed an expert system shell called *Negoplan* to support single party participants in a negotiation. *Negoplan* simulates the changes in the positions of the parties during the negotiation, based on their anticipated behavior. Their method does not simulate the entire process of negotiation since they give one party a competitive advantage. In simulating the overall simulation process, we concentrate on comparisons between one attribute subject of the negotiation and the outside options available to the negotiator.

In the work of Kraus, Lehmann and Ephrati, [26, 24, 27, 25] a general structure for a negotiator-agent was developed that functions in a complex environment, and several techniques for the performance of different tasks by such an agent were also developed. In the present study, we take a similar approach, while attempting to model a real world situation. In addition, we want to concentrate on a somewhat simplified case —less players, less issues to negotiate about— in order to be able to isolate different aspect of negotiations in such environments, to develop general theorems, and subject them to testing with computer models and human players.

2.2 Related Work in Economics and Game Theory

There are two main approaches for the development of theorems relating to the negotiation process. The first is informal theories which attempt to identify possible strategies for a negotiator and to assist a negotiator in achieving optimal results (see [13, 8, 22]). The other approach is the formal theory of bargaining originating with the work of John Nash ([43] [44]), who attempted to construct formal models of negotiation environments and to prove different theorems about the best strategies a negotiator can follow under different circumstances. This formal game theory approach provides clear analyses of various situations and precise results concerning the strategy a negotiator should choose. On the other hand, it requires making restrictive assumptions that are unacceptable to the first group.

Following Genesereth, Ginsberg, Rosenschein and Doyle, [16, 6, 7], we propose the use of game-theoretic techniques for Artificial Intelligence purposes. We propose to develop a strategic model of negotiation that can serve as the basis for building efficient automated negotiators. We realize that some of the assumptions we will be forced to make in developing the general strategic model will not be applicable in some situations, and in such cases we intend to use the informal theorems, referred to above, in order to fill in the gaps (in this respect our approach is similar to Raiffa [46]).

The formal game theory approach is also divided into two central sub-approaches concerning the bargaining problem (see [21]). The first is the strategic approach. The players' negotiating maneuvers are moves in a noncooperative game and the rationality assumption is expressed by investigation of the *Nash Equilibrium*.[1]

The second approach is the axiomatic method. It makes assumptions about the solution of a negotiation situation without specifying the bargaining process itself (the literature on the axiomatic approach to bargaining is surveyed by Roth [48] ; [40] is a good introduction to game theory).

Since we intend to use our theoretical work as a basis for the development of automated negotiators, we have adopted the strategic approach. Rubinstein [49] and Ståhl [57] developed models of alternating offers, which take time into consideration. Shaked and Sutton [54] extended these works by developing models in which a player can opt out of the game. Those works are closely related to our desired models (see [45] for a detailed review of the bargaining game of alternating offers). Nevertheless, several important modification are needed. These mainly concern the way time influences the preferences of the agents, the possibility that both agents can opt out, and the preferences of the agents over opting out.

2.3 Related Work in Crisis Analysis

Decision theorists have dealt quite extensively with the development of negotiation and bargaining strategy (for an excellent review of this literature, see [60]). The analysis of negotiation and bargaining behavior in crisis situations has fallen predominantly in the domain of political science. Studies in this area include a focus on deterrence [18], the bargaining process itself [56], cross national models of crisis decision making [2, 58], cognitive closure and crisis management [36], quantitative analysis of bargaining [38] and studies of crisis prevention [17]. Comprehensive statistical analysis of the behavior of states in crises is reported

[1] A pair of strategies (σ, τ) is a Nash Equilibrium if, given τ, no strategy of Player 1 results in an outcome that Player 1 prefers to the outcome generated by (σ, τ) and similarly for Player 2 given σ.

in Brecher et al. [4], Wilkenfeld et al. ([62], and Brecher and Wilkenfeld [3].

Our approach in the development of a model of strategic negotiation has been guided most directly by two studies. According to Snyder and Diesing [56], the three types of bargaining in crisis are *accommodative*, *coercive*, and *persuasive.* In the accommodative approach, we note a convergence of the bargaining positions of the parties toward a settlement through a sequence of bids or proposals for settlement, involving demands, offers, and concessions. Coercive bargaining is a process of showing firmness, involving threats and warnings, and in general exerting pressure to influence the other party to accept one's position. Coercion includes the threat of harm. Persuasion also attempts to influence the other party to accept one's position, but does not involve threatening harm. Both coercive and accommodative moves (threats and concessions) present the adversary with a choice between a pair of outcomes, one certain and the other uncertain. Persuasion involves moving the choice to one's own advantage [56][195-198].

A second typology with relevance to behavior patterns in crisis bargaining is proposed by Leng ([37]; see also [39].) Unlike Snyder and Diesing, Leng's typology is based on the joint behavior of the crisis dyad. Among the relevant behaviors examined are: *Fight*: The antagonists employ mutually coercive influence strategies, with the level of conflict spiraling upward to very high levels of conflictive behavior (1967 Six Day War); *Resistance*: One antagonist pursues a coercive strategy while the other stands firm. This produces a relatively moderate rate of escalation (Italy-Ethiopia 1935); *Standoff*: Both parties demonstrate firmness through threats, and neither is willing to retreat from its stand, but neither is willing to increase the level of tension beyond a certain point. This usually ends in compromise or stalemate (Berlin Wall 1961); *Dialogue*: Both sides pursue accommodative bargaining strategies. Escalation is low and reciprocity is high (Morocco Crisis 1905-6); and *Prudence*: One party is assertive, leading to rapid submission by the other (Austria Anschluss 1938) ([37][182-194]).

While none of these approaches is directly incorporated into the strategic model of negotiation presented below, they have helped sharpen our conception of the process and helped us distill its central elements.

Another related work is of Fraser and Hipel [14]. They developed a formal method that permits a rapid assessment of complex conflict situations for the purpose of finding resolution to a conflict. The output from the analysis includes possible stable solutions to the conflict. Comparing their work to ours, we model the process of the negotiation itself, taking into account the passage of time during the negotiation. Our analysis provides negotiation strategies for the players that are in perfect equilibrium.

3 The Hostage Crisis

The specific scenario which evolved during the course of our formalization of the crisis negotiation model was based on the hypothetical hijacking of a commercial airliner en route from Europe to Israel and its forced landing at Cairo International Airport. The passengers are predominantly Israeli, but there are a number of other nationals aboard. The hijackers are known to be Palestinian, although their precise affiliation is not immediately clear (and hence the credibility of their threats is not known at the outset). The hijackers will eventually demand the release from Israeli security prisons of an undetermined number of Arab prisoners, and safe passage for the hijackers to an as yet undisclosed destination (for additional details see [32].)

The hostage crisis was chosen as a typical case of multiparty negotiation. Although this hypothetical case is quite specific in details, the intention is to build a general model of negotiation. The choice of a real historical case would have increased the complexity of the model while at the same time reducing its potential generalizability.

Once the case was chosen, it was reduced to its essential characteristics. For example, this model consists of only three players: the terrorists, Israel, and Egypt (the latter plays the role of third party or mediator). We could have added additional players like the US or Syria, but we feel that these three adequately represent the most important types of players and their interests in such a negotiation. Similarly, we could have increased the number of options available to each player – for example, Israel could have had the option of kidnapping a prominent Palestinian leader, in addition to its two options of agreement with the terrorists or launching a military operation. Here again, we assume that the added complexity which additional options would entail would not add appreciably to the reliability or generalizability of the model.

Israel, the terrorists (hijackers), and Egypt must consider six possible outcomes:

1. Israel launches a military operation to free the hostages
2. Egypt launches a military operation to free the hostages
3. The terrorists blow up the plane with all aboard
4. Israel and the terrorists negotiate a deal involving the release of prisoners in Israeli jails, release of hostages, and safe passage for the terrorists
5. Egypt and the terrorists negotiate a deal involving release of the hostages and safe passage for the terrorists
6. The terrorists give up.

Each party to the negotiation has a set of objectives, and a certain number of utility points is associated with each (see [28]). Utility points were assigned in order to express a complex set of preferences in such a way that subtle distinctions can be made among them. Short term objectives pertain to the resolution or management of the immediate crisis, while long term objectives have to do with the consequences for the policy of that actor once the immediate situation has been resolved.

For Israel, short-term objectives involve the safe return of the passengers and an acceptable level of casualties among Israeli military personnel in the event of military action. For the terrorists, short-term objectives include the release of prisoners held in Israeli jails, release of the hostages, and safe passage for the terrorists. Egypt is cast in the role of mediator or facilitator, and has no exclusively short term goals.

Among Israel's major long-term goals is a cluster of factors relating to the credibility of its deterrence against terrorism, its overall strategic interests, and experience in counterterrorism. For the terrorists, long-term objectives include damage to Israel's internal and external image, damage to Israel's deterrence against terrorism, and damage to Israel's relations with the US and Egypt. For both Israel and the terrorists, the long term consequences are considerably more important than the resolution of the immediate situation.

As we have indicated, all of Egypt's objectives are long-term in nature. By far the most important Egyptian objective is its ability to demonstrate its control of the situation, and the maintenance of its internal image. Also of critical importance is Egypt's ability to emerge from the crisis with its relations with other Arab countries intact.

In combining the range of utility points associated with each objective with the six possible outcomes listed above, a matrix is generated which yields a point output total for the various outcomes. In the case of three of these outcomes – an Israeli or Egyptian military operation, and a terrorist decision to blow up the plane – probabilities are attached to the success or failure of such actions.

The specific issues to be negotiated during the course of the crisis include the following:

1. Israel-Terrorists
   (a) Number of prisoners to be released by Israel in exchange for release of the hostages
2. Israel-Egypt
   (a) Israel request for logistical information from Egypt on location and condition of plane, number and affiliation of hijackers and types of arms possessed by hijackers.
   (b) Israel request for Egyptian assistance during an Israeli operation
   (c) Israel request that Egypt deny the terrorists access to the media in order to publicize their message
   (d) Egypt request for Israeli assistance during an Egyptian operation
   (e) Egypt request that Israel accept a terrorist offer
3. Terrorists-Egypt
   (a) Terrorist request for access to the media to publicize their message
   (b) Egypt request that terrorist give up or reach an agreement for safe passage.
   (c) Egypt request that the terrorists accept an Israeli offer

The concept of the passage of time is incorporated into the model in two ways. First, it provides a reference point for the calculation of utilities and probabilities. Second, time is a factor for the three parties, since the passage of time impacts on them differentially.

In general, time works in favor of the terrorists, and against Israel and Egypt. This latter aspect of time sets up a complex negotiation dynamic for the crisis.

In general, time impacts on the following aspects of the model: (1) the probability of success of an Israeli or Egyptian military operation (having to do with whether the operation is launched in daylight or at night, time available for preparation of troops, deteriorating weather conditions, and condition of terrorists and hostages); (2) the extent of publicity for the terrorists' message; and (3) Israel and Egypt's internal and external images.

In the next section we will suggest a general negotiation model that can be used to capture some important properties of the negotiations taking place among the parties under the conditions outlined above.[2]

4 The Strategic Model of Negotiation

In this section we will describe a strategic model of negotiation. Any strategic model includes a detailed description of a bargaining procedure. Ours is a modification of Rubinstein's model of alternative offers which focuses on the passage of time and the preferences of the players for different agreements as well as for opting out of the negotiations [49].

The outcomes of the model will be defined as perfect equilibria which require that a player's strategy be optimal in each step of the game.

Using these notions we will analyze different kinds of negotiation situations. We will concentrate on cases where one of the players gains over time and the other loses (at least up to some period of time).

4.1 Description

We assume that the negotiation process is taking place during a crisis, where two players, the "Initiator" (I) of the crisis (terrorists) and the "Participant (against his will)" (P) in the crisis (Israel), are bargaining about the partition of M units of a desirable object (800 security prisoners in Israeli jails). The partitioning takes place only after both players have reached an agreement. In this model we focus on the negotiation process between Israel and the terrorists, and assume that Egyptian behavior is fixed and known.

[2]We focus in this paper on the bilateral case. See [31] for an extension of the model to an n-player game.

Negotiation is an iterative process that may include several iterations and may even continue forever. We assume that agents can take actions only at certain times in the set $\mathcal{T} = \{0, 1, 2...\}$.

In each period $t \in \mathcal{T}$ one agent, say i, proposes an agreement, and the other agent (j) either accepts the offer (Y) or rejects it (N) or opts out of the negotiation (O).[3] If the offer is accepted, then the negotiation ends, and the agreement is implemented. Also, opting out by j ends the negotiation. After a rejection, the rejecting agent then has to 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.

We will now present formal definitions pertaining to the negotiation structure.

Definition 4.1 Agreement:
An agreement is a pair (s_I, s_P), in which s_i is agent i's portion of the desirable object. The set of possible agreements is

$$S = \{(s_I, s_P) \in \mathbb{N}^2 : s_I + s_P = M\}$$

Throughout the rest of the paper, I's portion in an agreement will be written first.

Definition 4.2 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 $\{Y, N, O\} \cup S$. We first define a strategy f for an agent i who is the first agent to make an offer. Let F be the set of all sequences of functions $f = \{f^t\}_{t=0}^{\infty}$, where $f^0 \in S$, for t even $f^t : S^{t-1} \rightarrow S$, and for t odd $f^t : S^t \rightarrow \{Y, N, O\}$ (S^t is the set of all sequences of length t of elements in S and Y, N and O are defined above). F is the set of all strategies of the player who starts the bargaining. Similarly, let G be the set of all strategies of the player who, in the first move, has to respond to the other player's offer; that is, G is the set of all sequences of functions $g = \{g^t\}_{t=0}^{\infty}$ such that for t even $g^t : S^t \rightarrow \{Y, N, O\}$ and for t odd $g^t : S^{t-1} \rightarrow S$.

We make no assumptions about who begins the negotiation process, i.e., who makes the first offer. If an agreement is never reached, and neither player opts out, we denote the outcome as "Disagreement" (D).

[3] In the Hostage Crisis situation, Israel can opt out by launching an army operation, and the terrorists can blow up the plane.

Let $\sigma(f,g)$ be a sequence of offers possibly ending with O in which player 1 (who can be either I or P) starts the bargaining and adopts $f \in F$, and player 2 adopts $g \in G$. Let $L(f,g)$ be the length of $\sigma(f,g)$ (where the length may be infinite). Let $La(f,g)$ be the last element of $\sigma(f,g)$ (if there is such an element). $La(f,g)$ may be either in S and in such a case we will call it the *partition* or may be O which denotes that one of the players opts out of the negotiation. We present a formal definition for the outcome of the negotiation, when the agents use the strategies f and g.

Definition 4.3 Outcome of the Negotiation:
The outcome function of the game is defined by

$$\mathcal{P}(f,g) = \begin{cases} D & \textit{if } L(f,g) = \infty \\ (La(f,g), L(f,g)-1) & \textit{otherwise} \end{cases}$$

Thus, the outcome (s,t) where $s \in S$ is interpreted as the reaching of agreement s in period t, (O,t) is interpreted as one of the players opting out of the negotiations, and the symbol D indicates a perpetual disagreement with no player opting out.

The last component of the model is the preference of the players on the set of outcomes. Each player has preferences for agreements reached at various points in time, and 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 player $i = I, P$ has a preference relation (complete, reflexive, and transitive) $\succeq_i$ on the set $\{S \times \mathcal{T}\} \cup \{\{O\} \times \mathcal{T}\} \cup \{D\}$.

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 making an offer that was rejected or rejecting an offer (see, for example, the discussion of "decision-regret" in [46]).

In studying the Hostage Crisis case we identified a set of conditions that the players' preference relations should satisfy. We determined that those conditions fit a wide variety of cases.

First we assume that the least-preferred outcome is disagreement (D).

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

The next two conditions, (A1) and (A2), concern the behavior of $\succ_i$ on $S \times \mathcal{T}$, i.e., agreements reached in different time periods. Condition (A1) requires that among agreements reached in the same period, Player i prefers larger numbers of units s_i.

A1 "Object" is desirable: If $r_i > s_i$, then $(r,t) \succ_i (s,t)$.

The next assumption greatly simplifies the structure of preferences among agreements. It requires that preferences between (s_1, t_1) and (s_2, t_2) depend only on s_1, s_2 and the differences between t_1 and t_2. Furthermore, we assume that the bargaining costs or gains are fixed.

A2 Agreement's Cost Over Time: Each player has a number c_i $i \in \{I, P\}$ such that: $(s, t_1) \succeq_i (\bar{s}, t_2)$ iff $(s_i + c_i * t_1) \geq (\bar{s}_i + c_i * t_2)$.
We assume that player I gains over time ($c_I > 0$) and that player P loses over time ($c_P < 0$), i.e., player P prefers to obtain any given number of units sooner rather than later, while player I prefers to obtain any given number of units later rather than sooner.[4]

We note that assumption (A2) does not hold for O and the preferences of the players for opting out in different periods of time do not change in a stationary way. Furthermore, the preferences of a player for opting out versus an agreement fluctuate across periods of time in a non-stationary fashion.[5] In the case of the Hostage Crisis this is due to different rates of change over time in the probabilities associated with success or failure of the actions taken when opting out.

We also assume that player P prefers to opt out sooner rather than later and vice versa for player I.

A3 Opting Out Over Time: If $t_1 < t_2$ $(O, t_1) \succ_P (O, t_2)$ and $(O, t_2) \succ_I (O, t_1)$.

[4]Previous work on models of alternating offers (see, for example [49], [54]) assumed that time is of value to all parties. The Hostage Crisis is a situation in which one side (the terrorists) gains over time, while the other side (Israel) loses over time. Another example of such a situation occurs when a company contests a government attempt to restrain its advertising of a harmful product – the longer the company can tie up the issue in court and continue to advertise, the more units it can sell; conversely, the longer the case drags on, the more the consuming public (the government's "client") will be harmed by continued consumption of the product.

[5]Shaked and Sutton [54] considered the case where the players' preferences for opting out versus an agreement changes in a stationary manner.

4.2 Perfect Equilibrium

A useful notion for finding a good strategy is the Nash Equilibrium ([44, 40]). If there is a unique equilibrium, and if it is known that a player is designed to use this strategy, no agent will prefer to use a strategy other than this one.

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 ([49]). Therefore, we will use the stronger notion of (subgame) perfect equilibrium (P.E.) (see [53]) which requires that the players' strategies induce an equilibrium in any subgame (see [29] for the full definition).

4.3 Zone of Possible Agreement

When analyzing the model, the main question is whether a possibility exists that the players 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. As we mentioned above, in our model the preferences of a player for opting out versus an agreement fluctuate across periods of time in a non-stationary fashion and there is no fixed $s \in S$ such that for every $t \in \mathcal{T}$, $(s,t) \sim (O,t)$ as in [54]. This is the result of our assumption that the utility function of opting out changes differently over time than the utility function of an agreement. Therefore, we need the following definition in order to compare agreements with opting out.

Definition 4.4 *For every $t \in \mathcal{T}$ and $i \in \{I,P\}$ let $Pos_i^t = \{s^t \mid (s^t,t) \succeq_i (O,t)\}$.*

If Pos_i^t is not empty we define $\hat{s}^{i,t} = \min_{\prec_i} Pos_i^t$, i.e., $\hat{s}^{i,t}$ is the worst agreement that can be reached in period t which is still better for agent i than opting out. In case such an agreement does not exist, we define $\hat{s}_I^{I,t} = M-1$ and $\hat{s}_I^{P,t} = -1$.

If Pos_i^t is not empty then there will be only one minimal $\hat{s}^{i,t}$; this is because of assumption **A1** above. In order to avoid a discussion of extreme cases we will make the following assumption: For every $t \in \mathcal{T}$ if $\hat{s}_I^{P,t} \geq 0$ then $(\hat{s}^{P,t},t) \succ_P (O,t)$ and $\hat{s}_I^{P,t} < M - c_I - 2$. It is easy to extend the results of this paper after relaxation of this assumptions, but it will make the proofs longer and less readable. We note that in our formalization of the hostage crisis these assumptions are valid. To make the notation easier we will also assume that for any $t \in \mathcal{T}$, $((-1, M+1), t) \succeq_P ((-1, M+1), t+1)$.

We now introduce two additional assumptions that will ensure that an agreement will be reached.

A4 Possible Agreement: For every $t \in \mathcal{T}$ $(\hat{s}^{P,t}, t) \succ_P (\hat{s}^{P,t+1}, t+1)$ and if $\hat{s}_I^{P,t} \geq 0$ then $(\hat{s}^{P,t}, t) \succ_I (O, t+1)$.

Assumption (A4) ensures that if there are some agreements player P prefers over opting out, then there is at least one of those agreements that player I also prefers over opting out in the next period. We note that the assumption $(\hat{s}^{P,t}, t) \succ_P (\hat{s}^{P,t+1}, t+1)$ is not derived from the assumption $(O, t) \succeq_P (O, t+1)$ (A3).

Assumption (A4) alone does not ensure that an agreement is always possible. Let us consider the case that player P prefers to opt out over any agreement in the first period, i.e., $\hat{s}^{P,0} = -1$. In this case, if I starts the negotiation, it will end immediately by P opting out. If P starts the negotiation, since it must make an offer, the crisis may end with an agreement.

Lemma 1 *Let $(\hat{f}, \hat{g})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A4. If $\hat{s}_I^{P,0} = -1$ and player I starts the negotiation then $P(\hat{f}, \hat{g}) = (O, 0)$. If player P starts the negotiation then if $(0,0) \succeq_I (O, 1)$, $P(\hat{f}, \hat{g}) = (0, 0)$ otherwise, $P(\hat{f}, \hat{g}) = (O, 1)$.*

Proof: The proofs of the lemmas and theorems appear in the appendix. ▮

We can conclude that another assumption is necessary to ensure that an agreement may be reached, which states that an agreement is possible at least in the first period.

A5 Possible Agreement in the First Period: $\hat{s}_I^{P,0} \geq 0$

$\hat{s}^{P,0} \in S$ is the worst agreement for player P in period 0 which is still better than opting out. So, the requirement that I's portion of this agreement will be at least zero, ensures that there exists at least one agreement player P prefers over opting out. Using assumption (A4) which ensures that if there are some agreements player P prefers over opting out, then there is at least one of those agreements that player I prefers over opting out in the next period; together with assumption (A3) which requires that player I prefers to opt out later rather than sooner, we may conclude that there also exists an agreement that player I prefers over opting out in the first period. We determined that assumptions (A4) and (A5) are valid in the Hostage Case, but we note that those assumptions do not necessarily mean that agreement will be reached in such situations.

We will show now that under the above assumptions, if there exists a period when player P will prefer opting out over any agreement and the game has not ended in prior periods, then an agreement will be reached in the period prior to this period. If it is P's turn, it will offer its preferable agreements among the agreement that I prefers over opting out in the next period, and visa versa if it is I's turn.

Lemma 2 *Let $(\hat{f}, \hat{g})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A5. Suppose for some $T \in \mathcal{T}$, $\hat{s}_I^{P,T} \geq 0$ and $\hat{s}_I^{P,T+1} = -1$. If it is P's turn then using his P.E. strategy he will suggest $max_{\succeq_P}\{s \mid (s,T) \succeq_I (O, T+1)\}$, and if it is I's turn he will suggest $\hat{s}^{P,T}$. In both cases the other party will accept the offer.*

4.4 Player P Loses More Than Player I Gains

In this section we will assume that player P's losses over time are greater than player I's gains. In this model, for any agreement in period $t \in \mathcal{T}$, there is no other agreement in the future that *both* players will prefer over this agreement. On the other hand if an agreement s in period t is small enough, one can find an agreement in a period earlier than t which both players prefer over s in period t. According to our assumptions, this property will cause the players to reach an agreement in the first period.

First we consider the case in which player P starts the negotiations. We prove that in each period if an agreement exists which player P prefers over opting out there exists such an agreement which player I cannot reject. The idea is the following. Player P will accept or make an offer only if it is better for him than opting out. If I receives an offer such that there is no better agreement for him in the future, and it is also better for P than opting out in the future, and if he prefers this offer over P opting out in the next period, he must accept this offer. Otherwise, if this agreement is rejected, P should opt out as soon as possible, since he cannot expect to do any better than opting out. But if I prefers the proposed agreement over P's opting out in the next time period, he should accept the offer. We will show that such an agreement really exists under our assumptions.

Lemma 3 *Let $(\hat{f}, \hat{g})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A5 such that $|c_P| \geq |c_I|+1$ and player 1 is of type P. For any $T \in \mathcal{T}$ such that T is even (player P's turn) and $\hat{s}^{P,T} \geq 0$ (agreement is still possible) there exists an x^T such that $(x^T, T) \succeq_P (O, T+1)$ and $\hat{g}^T(x^T) = Y$.*[6]

We would like to define the x^T which is preferred by player P. That is, the best agreement for P that is acceptable to I.

Definition 4.5 *Let $(\hat{f}, \hat{g})$ be a Perfect Equilibrium (P.E.) as in lemma 3. For every $T \in \mathcal{T}$ we denote by $\hat{x}^T$ the maximal agreement (with regard to $\succeq_P$,) such that $(x^T, T) \succeq_P (O, T+1)$ and $\hat{g}^T(x^T) = Y$.*

The next lemma claims that the value of $\hat{x}^T$ depends only on $\hat{s}^{P,T+1}$ and on c_I.

Lemma 4 *If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} \leq c_I$ then for every $T \in \mathcal{T}$, $\hat{x}^T$ which is defined in definition 4.5, is equal to $(\hat{s}_I^{P,T+1} + 1 + c_I, \hat{s}_P^{P,T+1} - 1 - c_I)$.*

We would like to prove a similar lemma to lemma 3 for player P, i.e. whether a suggestion exists which player P will always accept. This is much easier since player P loses over time and therefore he will always accept an agreement that is better for him than the best agreement he can reach in the next period (i.e., $(0, M)$), which is also better for him than opting out.

Lemma 5 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that P is the first player. For any $T \in \mathcal{T}$ such that T is odd (player I's turn) and $\hat{s}_I^{P,T} \geq 0$ (agreement is still possible) if $s_I^T < min\{\hat{s}_I^{P,T}, -c_P\}$ then $f^T(s^T) = Y$.*

We will show now that any agreement that will be reached in some period $T \in \mathcal{T}$ where there is still a possibility for reaching an agreement in the next time period, will be at most

[6]For any $T \in \mathcal{T}$, such that T is even if for any $s_0, ..., s_{T-1}$ $f(s_0, ..., s_{T-1})$ does not depend on $s_0, ..., s_{T-1}$ we will denote the result by f^T and if for any $s \in S$ $g(s_0, ..., s_{T-1}, s)$ does not depend on $s_0, ..., s_{T-1}$, we will denote the result by g^T. Similarly, when T is odd.

(from P's point of view) the worst agreement to P which is still better to it than opting out, i.e., $\hat{s}^{P,T}$. The reason for that is that if there is still a possibility for an agreement in the next period, I wants to delay reaching an agreement. By offering $\hat{s}^{P,T}$ he prevents P from opting out, and gains another period of time. On the other hand, I won't accept anything worth less to him than $\hat{s}^{P,T}$, since he can always wait until the next period, gain a period, and reach such an agreement.

Lemma 6 *Let $(\hat{g}, \hat{f})$ be a P.E. of a model satisfying A0-A5. If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} \leq c_I$, $|c_P| \geq c_I + 1$, and $s_I^{P,T+1} \geq 0$ then if it is I's turn $\hat{g}_I^T \geq \hat{s}_I^{P,T}$ and if it is P's turn and $\hat{g}^T(s) = Y$ then $s_I \geq \hat{s}_I^{P,T}$.*

We will show in the next lemma that player I won't offer less in period $T \in \mathcal{T}$ than $\hat{s}^{P,T}$. This is mainly since otherwise, P will opt out.

Lemma 7 *Let $(\hat{g}, \hat{f})$ be a P.E. of a model satisfying A0-A5. If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} \leq c_I$, $|c_P| \geq c_I + 1$, and $\hat{s}_I^{P,T+1} \geq 0$ then if it is I's turn $\hat{g}_I^T \leq \hat{s}_I^{P,T}$.*

We would like to prove now that player P does not have a better strategy than to offer $\hat{x}^{P,T}$ whenever it is his turn to make an offer and he still prefers an agreement over opting out. This is since P may receive in the future (some period t), at most $\hat{s}^{P,t}$. But $\hat{x}^{P,T}$ which is the best agreement for P which is acceptable to I in period T is better for P that $\hat{s}^{P,t}$ in the future.

Lemma 8 *Let $(\hat{g}, \hat{f})$ be a P.E. of a model satisfying A0-A5. If $\hat{s}^{P,t} - \hat{s}^{P,t+1} \leq c_I$, $|c_P| \geq c_I + 1$, and $s_I^{P,T+1} \geq 0$ then if it is P's turn, $\hat{f}^T = \hat{x}^T$.*

The next theorem summarizes our results.

Theorem 1 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that Pstarts the bargaining. If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t-1} \leq c_I$ and $|c_P| \geq c_I + 1$, then $P(\hat{f}, \hat{g}) = ((\hat{s}_I^{P,1} + 1 + c_I, \hat{s}_P^{P,1} - 1 - c_I, 0)$*

Up to this point we have assumed that player P starts the bargaining. Let us consider now what happens when player I starts the bargaining. One can notice that $\hat{x}_I^T$ may be greater than $\hat{s}_I^{P,T}$. since $\hat{s}_I^{P,T} - \hat{s}_I^{P,T+1} \leq c_I$. Therefore, a suggestion of $\hat{x}^0$ won't be accepted by player P in the first period. On the other hand, $\hat{s}^{P,0}$ is acceptable to P, since it can't get anything better in the future and it is better than opting out.

Theorem 2 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that the first player is of type I. If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t-1} \leq c_I$ and $|c_P| \geq c_I + 1$, then $P(\hat{g}, \hat{f}) = (\hat{s}^{P,0}, 0)$*

4.5 Player I Gains More Than Player P Loses

Up to this point we have assumed that $|c_P| \geq c_I + 1$ (this is the case in the terrorist-Israel negotiations). We now want to consider the case when $|c_P| < c_I + 1$.[7] In such a case for any suggestion, if it is big enough, it is possible to find a suggestion in the future that will be better for both sides. Although it might appear that such an assumption will cause long delays in reaching an agreement, we will prove that in fact the delay will be at most one period.

The intuition behind this proof is as follows. If it is not player P's turn to make an offer in some time period t, he can always opt out and gain something equivalent to $\hat{s}^{P,t}$. So, in time period $t-1$ P will never make a better offer to I than $\hat{s}_I^{P,t} - c_P$, which is his benefit from opting out in the next period with the addition of P's loss over time. But I will refuse such an offer, since I prefers waiting a period and offering P $\hat{s}^{P,t} \in S$. This offer will prevent P from opting out, and if he accepts the offer I's share will be $\hat{s}_I^{P,t} + c_I$ which is better to I than $\hat{s}_I^{P,t} - c_P$ since $|c_P| < c_I + 1$.

First we will show that an agreement won't be achieved when it is P's turn to make an offer and there is still the possibility of an agreement in the next period.

Lemma 9 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| < c_I + 1$ and P is the first player. If it is P's turn in time period $T \in \mathcal{T}$ and $\hat{s}_I^{P,T+1} \geq 0$ then $f_I^T \leq \hat{s}_I^{P,T+1} - c_P$ and $\hat{g}^T = N$.*

In the next lemma we will prove that in any period t when it is player I's turn he will offer at most (from P's point of view) $\hat{s}^{P,t}$. It will prevent P from opting out.

Lemma 10 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| < c_I + 1$ and P is the first player. If it is I's turn and $\hat{s}_I^{P,T+1} \geq 0$ then $\hat{g}_I^{P,T} \geq \hat{s}^{P,T}$.*

---

[7] This is the situation when autonomous agents negotiate over the sharing of a common resource, and one of them is using the resource during the negotiation ([30]).

In the next lemma we show that if I offers P something less preferable by P than $\hat{s}^{P,t}$, P will opt out.

Lemma 11 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| \leq c_I + 1$ and P is the first player. If T is odd (i.e., I's turn) and $\hat{s}_I^{P,T+1} \geq 0$ then if $\hat{g}_I^T > \hat{s}_I^{P,T}$ $\hat{f}^T = O$*

The next theorem summarizes our results in this case. It claims that under our assumptions an agreement will be reached in period 1.

Theorem 3 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| < c_I + 1$, then if P is the first player $P(\hat{f}, \hat{g}) = (\hat{s}^{P,1}, 1)$.*

We will now prove a similar theorem for the case in which player I is the first player.

Theorem 4 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| < c_I + 1$, then if I is the first player $P(\hat{f}, \hat{g}) = (\hat{s}^{P,0}, 0)$.*

4.6 Player I's situation changes from winning to losing over time

In the hostage crisis there is some point when the terrorists stop gaining and start losing over time (see [28]). We want to consider such models and to show that our results are still valid.

Suppose there is some $T_c \geq 2$ such that from that period on player I stops gaining and starts losing over time. We assume that I's loses after T_c are smaller than P's loses at these time periods. That still gives I an advantage over P.

Formally we would like to replace (A2) by (A2').

A2' Agreement's Cost Over Time: Let $T_c, t_1, t_2 \in \mathcal{T}$ such that $T_c \geq 2$. We will assume that $t_i = t_i^1 + t_i^2$ where if $t_i \geq T_c$ $t_i^1 = T_c$ otherwise then $t_i^2 = 0$. $(s, t_1) \succeq_i (\bar{s}, t_2)$ iff $(s_i + c_i * t_1^1 + c_i' * t_1^2) \geq (\bar{s}_i + c_i * t_2^1 + c_i' * t_2^2)$ where $c_i' < 0$ $c_P < 0$ $c_I > 0$ $|c_I'| < |c_P'|$ and $|c_P'| < |c_P|$

Furthermore, we assume that after T_c I prefers to opt out sooner rather than later, but until T_c he prefers to opt out later rather than sooner. We don't make any assumptions concerning I's preference between opting out before T_c and opting out after this time period.

Case	First Player	$\hat{s}_I^{P,0} \geq 0$	$\hat{s}_I^{P,1} \geq 0$	$\|c_P\| \geq c_I + 1$	$\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} \leq c_I$	Result $\mathcal{P}(\hat{f}, \hat{g})$
1	P	yes	yes	yes	yes	$(\hat{s}_I^{P,1} + c_I + 1, 0)$
2	I	yes	—	—	—	$(\hat{s}_I^{P,0}, 0)$
3	I	no	no	—	—	$(O, 0)$
4	P	no	no	—	—	If $(0,0) \succeq_I (O,1)$, $(0,0)$ else $(O,0)$
5	P	yes	no	—	—	$(min\{s \| (s,0) \succ_I (O,1)\}, 0)$
6	P	yes	yes	no	yes	$(\hat{s}_I^{P,1}, 1)$

Figure 1: Summary of results. c_P is the loss of player P over time and c_I is the gain of player I over time. $\hat{s}^{P,t}$ is the worst agreement for player P in period t which is still better than opting out. — indicates that the result does not depend on the value of this field. Each agreement in the result's column is denoted by I's portion in that agreement. For example $(\hat{s}_I^{0,P}, 0)$ is actually $((\hat{s}_I^{0,P}, M - \hat{s}_I^{0,P}), 0)$

Formally, we replace (A3) by (A3').

A3' Opting Out Over Time: If $t_1 < t_2$ $(O, t_1) \succ_P (O, t_2)$ and if $t_1 < t_2 < T_c$ $(O, t_2) \succ_I (O, t_1)$ and if $t_1 > T_c$, $(O, t_1) \succ_I (O, t_2)$.

We note that theorem 1 is still valid. This is mainly because even though player I starts losing from period T_c, our assumption that his losses are smaller than those of player P leaves him in a position which is still better than player P. Therefore, what player P can gain from the subgame starting in period T_c is not better than $(\hat{x}^{T_c}, 0)$ if he starts the game or $(\hat{s}^{P,T_c}, 0)$ if player I starts the game.

4.7 Factors Influencing the Outcome

Figure 1 presents a general summary of our results concerning the model's behavior under different conditions. We focus on the identification of the most important factors which influence the outcomes of negotiations under differing circumstances. For example, in Case 1 player P makes the first offer. The "yes"s in the third and fourth columns indicate that the worst agreement to player P which he still prefers over opting out is not less than zero, i.e., there exists an agreement in those periods which he prefers over opting out. The "yes" in the fifth column indicates that player P loses over time more than player I gains. The last column indicates that an agreement will be reached in the first period and player I will get $\hat{s}_I^{P,1} + c_I + 1$ and player P $M - (\hat{s}_I^{P,1} + c_I + 1)$. The fourth column in Case 2, for example, indicates that this case actually represents multiple cases, i.e., $\hat{s}^{P,1}$ can be either at least zero, but also may be less than zero.

The single most important influence on the nature of the agreement reached by the players is $\hat{s}^{P,t}$, i.e., the worst agreement which player P will agree to in period t which is still better for him than opting out in this period (see cases 1,2,6 in Figure 1). The intuition behind this result is that since we assume that player I also prefers an agreement of $\hat{s}^{P,t}$ over the option of opting out in period t, player P's threat of opting out turns out to be credible. On the other hand, since player I gains over time, he can afford to wait and therefore he does not have to suggest or accept an offer which is worse for him than $\hat{s}^{P,t}$. From the property that $\hat{s}_I^{I,t} \leq \hat{s}_I^{P,t}$ (by A4) and by reasoning similar to the proceeding, $\hat{s}^{I,t}$ does not influence the outcome of an agreement.

A second factor which influences the outcome is the question of which player begins the negotiation process. If player P starts the negotiation, it creates an advantage for player I (compare cases 1 and 2, cases 6 and 2 and 3 and 4). This is the case because it delays the threat of player P to opt out for at least one period. Since the passage of time works to the advantage of player I and against player P, player I benefits from the delay. However, if we assume that all players are allowed to opt out in any period, the result will change, since I will then lose part of its advantage.

The ratio of how much player I gains over time (c_I) to how much player P loses over time (c_P) determines the time period in which an agreement is reached, when player P starts the negotiation (compare cases 1 and 6 in Figure 1). If $c_P \geq c_I + 1$, then agreement will be reached in the first period (if in fact an agreement is reached at all), and if $c_P < c_I + 1$, then an agreement will be reached in the second period. This results from a property of the second model, that no agreement in a current period can be better for both parties than a specific agreement in the future. In particular, no agreement in period t is better for both parties than reaching an agreement $\hat{s}^{P,t+1}$ in period $t + 1$.

5 The Hostage Crisis Revisited

Since one of our aims in developing the above theorems is to use them in the specific negotiation situation we deal with, we want to demonstrate how these theorems are applicable in this case. We concentrate on the Israel-terrorist negotiation process which involves the number of prisoners that will be released from Israeli prisons in exchange for the release of all the hostages on the plane.

In order for the results of Section 4 will be applicable, we need to make the following assumptions:

1. Full information - Israel and the terrorists know each others preferences.
2. Rational behavior - both sides use the notion of perfect equilibrium when choosing their strategies.
3. Egypt's behavior is fixed and known to both sides.
4. Commitments are enforced.
5. Assumptions (1)-(4) are common knowledge.

In our case $S = \{(0,799),(1,798),...,(799,0)\}$. Since we assume that an agreement between Israel and the terrorists requires that Israel release at least one prisoner [28], an agreement $(s_1, s_2) \in S$ is interpreted as Israel releasing $s_1 + 1$ prisoners. Israel is a player of type P and the terrorists initiate the crisis (type I). We denote Israel by P_{Is} and the terrorists by I_T.

As we explained in Section 3, while formalizing the situation we attached utilities and probabilities to the possible outcomes and the way they are changing over time. We determined that through time period 10, the terrorists gain over time, and from time period 11 on, the terrorists lose over time; Israel loses over time across all periods of the game.

We determined that Israel's utility from releasing x prisoners depends on three factors: a positive constant ($D_{P_{I_s}}$) which is determined by Egypt's behavior, a constant loss from the release of each prisoner ($v_{P_{I_s}}$) and the loss over time $d^i_{P_{I_s}}$. The loss over time changes after period 10.

So, formally, the outcome for Israel from releasing x prisoners in period $t \leq 10$ is $D_{P_{Is}} + x * v_{P_{Is}} + d^1_{P_{Is}} * t$. If $t > 10$ then the outcome for Israel from releasing x prisoners is $D_{P_{Is}} - x * v_{P_{Is}} + d^1_{P_{Is}} * 10 + d^2_{P_{Is}} * (t-10)$, where $d^1_{P_{Is}} < d^2_{P_{Is}} < 0$.

Similarly, the utility of the terrorists from Israel's release of x prisoners also depends on three factors. A positive constant (D_{I_T}) which is determined by Egypt's behavior, a constant gain from the releasing of each prisoner (v_{I_T}) and the gain until time period 10 ($d^1_{I_T}$) and loss over time after time period 10 ($d^2_{I_T}$).

So, formally, the outcome for the terrorists from Israel releasing x prisoners in period t, while Egypt's behavior is known, is if $t < 10$ $D_{I_T} + x * v_{I_T} + d^1_{I_T} * t$ where $d^1_{I_T} > 0$ and if $t \geq 10$, $D_{I_T} + x * v_{I_T} + d^1_{I_T} * 10 + d^2_{I_T} * (t-10)$ where $d^1_{I_T} > 0$ and $d^2_{I_T} < 0$.

Now, using those utilities one can compute the preferences of both Israel and the terrorists for possible agreements. c_i where $i \in \{P_{Is}, I_T\}$ (see (A2) in Section 4.1) is computable as follows: $c_i = d_i / v_i$.

It is more difficult to analyze the preferences of the parties for opting out in different periods and especially difficult to compare the option of opting out with an agreement. Nevertheless, we were able to establish that assumptions of section 4 hold in our case;

A0 Disagreement is the worst outcome.

A1 The prisoners are desirable.

A2 Agreement's Cost Over Time– as we demonstrated above, Israel prefers to release any given number of prisoners sooner rather than later, while the terrorists through period 10 prefer to obtain any given number of prisoners later rather than sooner.

A3 Opting Out Over Time — Israel prefers to opt out sooner rather than later.[8] The terrorists, through period 10, prefer opting out later rather than sooner.

A4 Possible Agreement — The terrorists prefer any possible agreement over opting out. Israel prefers its worst agreement in time period t which is still better than opting out,

[8]Opting out in the Hostage Crisis situation does not mean opting out immediately, but rather making that decision. Actually, Israel prefers to opt out in time period 3, but makes the decision in the first time period.

in time t $(\hat{s}^{P_{Is},t}, t)$ over the worst agreement in time period $t+1$ which is still better than opting out, in time $t+1$ $(\hat{s}^{P_{Is},t+1}, t+1)$.

A5 Possible Agreement in the First Period.

Now, we need to determine which of the conditions discussed in Section 4 actually occur in our case. That is, we need to check how the factors that influence the outcomes operate in the hostage crisis.

In the case of any Egyptian behavior $|c_{P_{Is}}| > c_{I_T} + 1$. Even though $\hat{s}^t_{P_{Is}} - \hat{s}^{t+1}_{P_{Is}}$ is not constant ($\hat{s}^t_{P_{Is}}$ is the worst agreement for Israel, which is of type P, in period t which is still better than opting out, see Definition 4.4), beside the case that Egypt does not give information to Israel but also does not help the terrorists, $\hat{s}^t_{P_{Is}} - \hat{s}^{t+1}_{P_{Is}} \leq c_{I_T}$.

Therefore we can use the results from Section 4.4. and conclude that under assumptions (1)-(5) above that agreement will be reached in the first period. The details of the agreement depend of who starts the negotiations. From our results one can also learn what strategies an automated Israeli or terrorist player needs to adopt if the simulation is played out under the above conditions. The authors are currently working on the design of such an automated negotiator.

6 Conclusion

This paper has presented a strategic model for negotiation of alternative offers which takes into account the effect of time on the negotiation process. In our model both players can opt out, and one of the players loses over time and the other gains over time (at least up to some future point). We show that if there is a zone for an agreement, an agreement will be reached in the first or the second period of the negotiation.

We formalized a hostage crisis and showed that under certain assumptions the strategic model can be used to analyze this crisis. One of the assumption was that Egypt's behavior is known and fixed throughout the negotiations. Even after relaxation of this assumption a player can use our results to determine his preferences for Egypt's behavior. This can be done by examining the factors that influence the outcomes in each of the cases.

Do we think that in real situations and even in simulations of such situations an agreement will be reached in the first or second periods? We suspect not (see [32, 20, 50]). Rather, we think that after relaxation of the full information condition and taking into consideration the behavior of a third party, delays in reaching an early agreement will appear. In future work we intend to develop models that will be able to capture these situations.

7 Acknowledgement

We would like to thank Kenneth Arrow, Robert Auman, Bruce Bueno de Mesquita, Joe Oppenheimer and Payton Young for helpful suggestions about different aspects of game theory in negotiations in general and in the hostage crisis in particular. We would also like to thank Yona Alexander, Doreen Bass, Michael Brecher, Raymond Cohen, Thomas Colosi, Michael Harris, and several officials of the Israeli Government for helpful comments and advice about formalizing the hostage crisis.

A Proofs of Theorems and Lemmas

Lemma 1 *Let $(\hat{f}, \hat{g})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A4. If $\hat{s}_I^{P,0} = -1$ and player I starts the negotiation then $P(\hat{f}, \hat{g}) = (O, 0)$. If player P starts the negotiation then if $(0,0) \succeq_I (O,1)$, $P(\hat{f}, \hat{g}) = (0,0)$ otherwise, $P(\hat{f}, \hat{g}) = (O,1)$.*

Proof: By (A4) and the assumption that $\hat{s}_I^{P,0} = -1$ it is clear that for every $t > 0$ $\hat{s}_I^{P,t} = -1$, i.e., there is no agreement that P prefers over opting out in any time period. By (A3) if $t_1 < t_2$ then $(O, t_1) \succeq_P (O, t_2)$, i.e., it is better for P to opt out sooner rather than later. Therefore, if I starts the negotiation P will opt out immediately.

If player P starts the negotiation he must make an offer.[9] By assumption (A2), $f^0 = 0$ and using similar arguments as in the previous case we can conclude that $f^1(s) = O$ for any $s \in S$. If $(0,0) \succ_I (O,1)$ then $\hat{g}^0(0) = Y$, otherwise, $\hat{g}^0(0) = N$. ∎

Lemma 2 *Let $(\hat{f}, \hat{g})$ be a Perfect Equilibrium (P.E.) of a model satisfying A0-A5. Suppose for some $T \in \mathcal{T}$, $\hat{s}_I^{P,T} \geq 0$ and $\hat{s}_I^{P,T+1} = -1$. If it is P's turn then using his P.E. strategy*

[9] This is an artificial situation since we allow the players to opt out only after rejecting an offer. In another paper we will consider the case that players can opt out in any period.

he will suggest $max_{\succeq_P}\{s\ |(s,T)\succeq_I (O,T+1)\}$*, and if it is I's turn he will suggest* $\hat{s}^{P,T}$*. In both cases the other party will accept the offer.*

Proof: Suppose that it is $P's$ turn. Since $\hat{s}_I^{P,T+1}=-1$ player P will always opt out in period $T+1$. By assumptions (A3), (A4) and by definition 4.4, $(max_{\succeq_P}\{s\ |(s,T)\succeq_I (O,T+1)\},T)\succeq_P (\hat{s}^{P,T},T)\succ_P (O,T)\succeq_P (O,T+1)$. By assumption (A3) $(max_{\succeq_P}\{s\ |(s,T)\succ_I (O,T+1)\},T)\succeq_I (O,T)$, and the claim is clear. Similar arguments hold when it is player I's turn. ∎

Lemma 3 *Let* $(\hat{f},\hat{g})$ *be a Perfect Equilibrium (P.E.) of a model satisfying A0-A5 such that* $|c_P|\geq|c_I|+1$ *and player 1 is of type P. For any* $T\in\mathcal{T}$ *such that T is even (player P's turn) and* $\hat{s}^{P,T}\geq 0$ *(agreement is still possible) there exists an* x^T *such that* $(x^T,T)\succeq_P (O,T+1)$ *and* $\hat{g}^T(x^T)=Y$*.*[10]

Proof: Let $(\hat{g},\hat{f})$ be a P.E. and let $T\in\mathcal{T}$ such that T is even and $\hat{s}^{P,T}\geq 0$.

We distinguish between two cases.

In the first case, $\hat{s}_I^{P,T+1}=-1$ (there is no possible agreement in the next period). By lemma 2 the claim is clear.

In the rest of the proof we assume that $\hat{s}^{P,T+1}\geq 0$.

Suppose there exists $x^T\in S$ satisfying:

1. For all $t\in\mathcal{T}$, $t>1$, if $(O,T+t)\succeq_I (x^T,T)$ then $(O,T+1)\succeq_P (O,T+t)$.
2. $(x^T,T)\succeq_P (O,T+1)$
3. $\forall x^i\in S$ such that for every $t\geq 1$ $(x^i,T+t)\succeq_I (x^T,T)$, $(O,T+1)\succ_P (x^i,T+t)$
4. $(x^T,T)\succ_I (O,T+1)$

then $\hat{g}^T(x^T)=Y$.

Suppose $\hat{g}^T(x^T)\neq Y$. Since $(O,t+1)\succ_I (O,t)$ [(A3)], by (4) it is clear that $\hat{g}^T(x^T)\neq O$. That is, it is better for I to let P opt out in the next period, than to opt out by himself in the current period.

[10]For any $T\in\mathcal{T}$, such that T is even if for any $s_0,...,s_{T-1}$ $f(s_0,...,s_{T-1})$ does not depend on $s_0,...,s_{T-1}$ we will denote the result by f^T and if for any $s\in S$ $g(s_0,...,s_{T-1},s)$ does not depend on $s_0,...,s_{T-1}$, we will denote the result by g^T. Similarly, when T is odd.

So, suppose $\hat{g}^T(x^T) = N$. We may consider two cases. If for some $t \geq 1$, $P(\hat{f}, \hat{g}) = (O, T+t)$ then $(O, T+t) \succeq_I (X^T, T)$. Since $(x^T, T) \succ_I (O, T+1)$ (by 4), we can conclude that $t \geq 2$. But $(O, T+1) \succ_P (O, T+t)$ by A4 and therefore if the crisis is ended by opting out, P will prefer to do it sooner rather than later, and not wait until $T+t$. So, we can conclude that $P(\hat{f}, \hat{g}), x^T) \neq (O, t)$.

Now suppose $P(\hat{f}, \hat{g}) = (s, T+t)$ for some $s \in S$ and $t \in \mathcal{T}$ and $t \geq 1$, such that $(s, T+t) \succeq_I (x^T, T)$. By assumption 3, $(O, T+1) \succ_P (s, T+t)$ and therefore $(\hat{f}, \hat{g})$ is not a Perfect Equilibrium. This is a contradiction and we can conclude that $\hat{g}^T(x^T) = Y$.

It is left to show that such an x^T exists.

1. Suppose there exists $t > 1$ such that $(O, T+t) \succeq_I (x^T, T)$. By (A3) we can conclude that $(O, T+1) \succeq_P (O, T+t)$.

2. If $(x^T, T) \succeq_P (O, T+1)$ iff $(x^T, T) \succeq_P (\hat{s}^{P,T+1}, T+1)$ (by definition 4.4) and by (A3) we get that $M - x_I^T + c_P * T \geq M - \hat{s}_I^{P,T+1} + c_P * (T+1)$ and $\hat{s}_I^{P,T+1} - c_P \geq x_I^T$
Now, since we assumed in the lemma that $\hat{s}_I^{P,T+1} \geq 0$ and since $c_P < 0$ there exists such $x^T \in S$.

3. Suppose $x^i \in S$ and $t \geq 1$. $(x^i, T+t) \succeq_I (x^T, T)$ iff $x_I^i + (T+t) * c_I \geq x_I^T + T * c_I$ by (A3) iff $x_I^i \geq x_I^T - t * c_I$ (*).
If $((\hat{s}_I^{P,T+1} + 1, \hat{s}_P^{P,T+1} - 1), T+1) \succeq_P (x^i, T+t)$ then $(O, T+1) \succ_P (x^i, T+t)$.
$((\hat{s}_I^{P,T+1} + 1, \hat{s}_P^{P,T+1} - 1), T+1) \succeq_P (x^i, T+t)$ holds iff $M - \hat{s}_I^{P,T+1} - 1 + c_P * (T+1) \geq M - x_I^i + (T+t) * c_P$ and we get $\hat{s}_I^{P,T+1} + 1 - c_P + t * c_P \leq x_I^i$ (**).

So, in order for all the x^i that satisfy (*) to also satisfy (**) $x_I^T - t * c_I \geq \hat{s}_I^{P,T+1} + 1 - c_P + t * c_P$ and we can conclude that $x_I^T \geq \hat{s}_I^{P,T+1} + 1 - c_P + t * (c_P + c_I)$.

4. By assumption (A4) $(\hat{s}^{P,T}, T) \succ_I (O, T+1)$ and therefore it is sufficient to show that $(x^T, T) \succeq_I (\hat{s}_P^T, T)$ and by (A3) that $x_I^T \geq \hat{s}_I^{P,T}$.

From 2,3 and 4 we get that $\hat{s}_I^{P,T+1} - c_P \geq x_I^T \geq max\{\hat{s}_I^{P,T+1} + 1 - c_P + t*(c_P + c_I), \hat{s}_I^{P,T}\}$. Now, by (A3) $c_P < 0$ and $c_I > 0$ and by (A4) $\hat{s}_I^{P,T} - \hat{s}_I^{P,T+1} < -c_P$. So, if $|c_P| \geq |c_I| + 1$ such an x^T exists. ∎

Lemma 4 *If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} \leq c_I$ then for every $T \in \mathcal{T}$, $\hat{x}^T$ which is defined in definition 4.5, is equal to $(\hat{s}_I^{P,T+1} + 1 + c_I, \hat{s}_P^{P,T+1} - 1 - c_I)$.*

Proof: By the proof of lemma 3, $x_I^T \geq max\{\hat{s}_I^{P,T+1} + 1 - c_P + t * (c_P + c_I), \hat{s}_I^{P,T}\}$. Since $\hat{s}_I^{P,T+1} + 1 - c_P + t * (c_P + c_I)$ is maximal when $t = 1$ we get that $x_I^T \geq max\{\hat{s}_I^{P,T+1} + 1 + c_I, \hat{s}_I^{P,T}\}$. Since we assumed that $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} < c_I$, $\hat{s}_I^{P,T+1} + c_I > \hat{s}_I^{P,T}$. We may conclude that $\hat{x}^T = (\hat{s}_I^{P,T+1} + 1 + c_I, \hat{s}_P^{P,T+1} - 1 - c_I) \in S$. ∎

Lemma 5 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that P is the first player. For any $T \in \mathcal{T}$ such that T is odd (player I's turn) and $\hat{s}_I^{P,T} \geq 0$ (agreement is still possible) if $s_I^T < min\{\hat{s}_I^{P,T}, -c_P\}$ then $f^T(s^T) = Y$.*

Proof: If $(s, T) \succ_P ((0, M), T + 1)$ and $(s, T) \succ_P (O, T)$ then by (A3) and since $(\hat{F}, \hat{g})$ are a P.E., $f^T(s) = Y$. But, from our assumptions, since $s_I < -c_P$, $(s, T) \succ_P ((0, 0), T + 1)$ and since $s_I \leq \hat{s}_I^{P,T}$, $(s, T) \succ_P (O, T)$. ∎

We will show now that any agreement that will be reached in some period $T \in \mathcal{T}$ where there is still a possibility for reaching an agreement in the next time period, will be at most (from P's point of view) the worst agreement to P which is still better to it than opting out, i.e., $\hat{s}^{P,T}$.

Lemma 6 *Let $(\hat{g}, \hat{f})$ be a P.E. of a model satisfying A0-A5. If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} \leq c_I$, $|c_P| \geq c_I + 1$, and $s_I^{P,T+1} \geq 0$ then if it is I's turn $\hat{g}_I^T \geq \hat{s}_I^{P,T}$ and if it is P's turn and $\hat{g}^T(s) = Y$ then $s_I \geq \hat{s}_I^{P,T}$.*

Proof: If there is still a possibility for an agreement in the next period, I wants to delay reaching an agreement. By offering $\hat{s}^{P,T}$ he prevents P from opting out, and gains another period of time.

On the other hand, I won't accept anything worth less to him than $\hat{s}^{P,T}$, since he can always wait until the next period, gain a period, and reach such an agreement. ∎

Lemma 7 *Let $(\hat{g}, \hat{f})$ be a P.E. of a model satisfying A0-A5. If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} \leq c_I$, $|c_P| \geq c_I + 1$, and $\hat{s}_I^{P,T+1} \geq 0$ then if it is I's turn $\hat{g}_I^T \leq \hat{s}_I^{P,T}$.*

Proof: Suppose $\hat{g}_I^T > \hat{s}^{P,T}$. By lemma 6, for every $t > T$ if $\hat{s}_I^{P,t+1} \geq 0$ any agreement that will be reached in any such period t will be at most $\hat{s}^{P,t}$.

But by definition 4.4, $(O,T) \succ_P (\hat{g}^T, T) \succ_P (\hat{g}^T, T)$. Furthermore, for any $t > T$ $(O,T) \succ_P (\hat{s}^{P,t}, t)$ since $(O,T) \succ_P (\hat{s}^{P,T}+1, T) \succeq (\hat{s}^{P,T+1}, T+1)$ and for every $t_1, t_2 \in \mathcal{T}$ if $t_1 \leq t_2$ then $(\hat{s}^{P,t_1}, t_1) \succ_P (\hat{s}^{P,t_2}, t_2)$.

In addition if $\hat{s}_I^{P,t+1} = -1$, $(O,T) \succ_P ((\hat{s}_I^{P,T}+1, \hat{s}_P^{P,T}-1), T) \succeq_P ((0,M)t)$. We may conclude that if $\hat{g}_I^T > \hat{s}^{P,T}$ P will opt out. But since by (A4) $(\hat{s}^{P,T}, T) \succ_I (O,T)$, the claim is clear. ∎

Lemma 8 *Let $(\hat{g}, \hat{f})$ be a P.E. of a model satisfying A0-A5. If $\hat{s}^{P,t} - \hat{s}^{P,t+1} \leq c_I$, $|c_P| \geq c_I + 1$, and $s_I^{P,T+1} \geq 0$ then if it is P's turn, $\hat{f}^T = \hat{x}^T$.*

Proof: By lemma 6, any agreement in period $t > T$ will be at most $\hat{s}^{P,t}$. But $(\hat{x}^T, T) \succ_P (\hat{s}^{P,t}, t)$ and $(\hat{x}^{P,T}, T) \succ_P (O, T+1)$ and by (A4) the claim is clear. ∎

Theorem 1 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that Pstarts the bargaining. If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t-1} \leq c_I$ and $|c_P| \geq c_I + 1$, then $P(\hat{f}, \hat{g}) = ((\hat{s}_I^{P,1} + 1 + c_I, \hat{s}_P^{P,1} - 1 - c_I, 0)$*

Proof: The proof is clear by the above lemmas. ∎

Theorem 2 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that the first player is of type I. If $\hat{s}_I^{P,t} - \hat{s}_I^{P,t-1} \leq c_I$ and $|c_P| \geq c_I + 1$, then $P(\hat{g}, \hat{f}) = (\hat{s}^{P,0}, 0)$*

Proof: Similar to lemmas 3 and 8, one can prove that, $g^1 = \hat{x}^1$ and that $f^1(x^1) = Y$. But, $(O,0) \succ_P (\hat{x}^1, 1)$ and therefore player I needs to suggest to Player P a better offer than opting out in the first period, i.e. $f^0 = \hat{s}^{P,0}$. Since $(\hat{s}^{P,0}, 0) \succ_P (\hat{x}^1, 1)$ the claim is clear. ∎

Lemma 9 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| < c_I + 1$ and P is the first player. If it is P's turn in time period $T \in \mathcal{T}$ and $\hat{s}_I^{P,T+1} \geq 0$ then $f_I^T \leq \hat{s}_I^{P,T+1} - c_P$ and $\hat{g}^T = N$.*

Proof: Any outcome of the crisis after time period T should be preferred by player P than $(O, T+1)$ since player P can always opt out in period $T+1$. For every $s \in S$ $(s, T) \succeq_P (O, T+1)$ iff $(s, T) \succeq_P (\hat{s}^{P,T+1}, T+1)$ iff $s_I \leq \hat{s}_I^{P,T+1} - c_P$. Therefore, $\hat{f}_I^T \leq \hat{s}_I^{P,T+1} - c_P$

It is clear that $(\hat{s}^{P,T+1}, T+1) \succeq_I ((\hat{s}_I^{P,T+1} - c_P, \hat{s}_P^{P,T+1} + c_P), T)$.

In order to show that $\hat{g}^T = N$ we need to distinguish between two cases.

1. Suppose there exists $T_0 \in \mathcal{T}$ such that $\hat{s}_I^{P,T_0} = -1$. Using lemma 1 and using similar arguments as in the proof of lemma 8 and since $c_I \geq c_P + 1$ one can conclude that any agreement that can be achieved in $T_0 - 1$ is better for player I than $((\hat{s}_I^{P,T+1} - c_P, \hat{s}_P^{P,T+1} + c_P), T)$. So, I can always offer P in time period $T < t < T_0$ $\hat{s}^{P,t}$, prevent it from opting out, and wait until $T_0 - 1$

2. If for every $t \in \mathcal{T}$, if $\hat{s}_I^{P,t+1} \geq 0$ then by A4 and since $c_I + 1 > |c_P|$ and we are dealing with a discrete field $(\hat{s}^{P,t+1}, t+1) \succeq_I (\hat{s}^{P,t}, t)$ and the claim is clear. ▮

In the next lemma we will prove that in any period t when it is player I's turn he will offer at least $\hat{s}^{P,t}$.

Lemma 10 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| < c_I + 1$ and P is the first player. If it is I's turn and $\hat{s}_I^{P,T+1} \geq 0$ then $\hat{g}_I^{P,T} \geq \hat{s}^{P,T}$.*

Proof: (Sketch) Since $c_I + 1 > |c_P|$, $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} \leq c_I$, and the proof is similar to the one in lemma 6. ▮

Lemma 11 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| \leq c_I + 1$ and P is the first player. If T is odd (i.e., I's turn) and $\hat{s}_I^{P,T+1} \geq 0$ then if $\hat{g}_I^T > \hat{s}_I^{P,T}$ $\hat{f}^T = O$*

Proof: (Sketch) From the assumption that $(\hat{s}^{P,t}, t) \succ_P (\hat{s}^{P,t+1}, t+1)$ and since we are dealing with a discrete case $\hat{s}_I^{P,t} - \hat{s}_I^{P,t+1} < -c_P - 1$. Therefore $(O, T) \succ_P ((\hat{s}_I^{P,T} + 1, \hat{s}_P^{P,T} - 1, T) \succeq_P (\hat{s}^{P,T+1}, T+1)$. But by (A4) and (A3) we can conclude that for every $t \in \mathcal{T}$, if $\hat{s}_I^{P,t} \geq 0$ $(\hat{s}^{P,T+1}, T+1) \succeq_P (\hat{s}^{P,T+t+1}, T+t+1)$. Therefore, it is sufficient to show that player I won't suggest or accept an offer which player P prefers over $(\hat{s}^{P,T+1}, T+1)$. But this is clear from lemma 10. ▮

Theorem 3 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| < c_I + 1$, then if P is the first player $P(\hat{f}, \hat{g}) = (\hat{s}^{P,1}, 1)$.*

Proof: By lemma 9 it is clear that an agreement won't be reached in the first period. By lemma 10 and lemma 11 and since $(\hat{s}^{P,1}, 1) \succ_I (O, 1)$ it is clear that $\hat{g}^1 = \hat{s}^{P,1}$. But by lemmas 9 and 10 and with similar arguments as in lemma 8 it is clear that for any $s \in S$, $\hat{f}^1(s, \hat{s}^{P,1}) = Y$. ▮

Theorem 4 *Let $(\hat{f}, \hat{g})$ be a P.E. of a model satisfying A0-A5 such that $|c_P| < c_I + 1$, then if I is the first player $P(\hat{f}, \hat{g}) = (\hat{s}^{P,0}, 0)$.*

Proof: It is clear that lemma 10 and lemma 11 are still valid under the assumption that I is the first player. Therefore $P(\hat{f}, \hat{g}) = (\hat{s}^{P,0}, 0)$.

References

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

[2] M. Brecher, editor. *Studies in Crisis Behavior.* Transaction Books, New Jersey, 1978.

[3] M. Brecher and J. Wilkenfeld. *Crisis, Conflict and Instability.* Pergamon Press, Oxford, 1989.

[4] M. Brecher, J. Wilkenfeld, and S. Moser. *Crises in the Twentieth Century: Vol.I Handbook of International Crises.* Pergamon Press, Oxford, 1988.

[5] 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.

[6] J. Doyle. Reasoning, representation, and rational self-government. In *Proc. of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 367–380, 1989.

[7] J. Doyle. Rationality and its role in reasoning. In *Proc. of AAAI-90*, Boston, MA, 1990. Invited Talk.

[8] D. Druckman. *Negotiations.* Sage Publications, 1977.

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

[10] 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.

[11] E. H. Durfee and T. A. Montgomery. A hierarchical protocol for coordinating multiagent behaviors. In *Proc. of AAAI-90*, pages 86–93, Boston, MA, 1990.

[12] R. E. Fikes, P.E. Hart, and N. J. Nilsson. Learning and executing generalized robot plans. *Artificial Intelligence*, 3(4):251–288, 1972.

[13] R. Fisher and W. Ury. *Getting To Yes: Negotiating Agreement Without Giving In.* Houghton Mifflin, Boston, 1981.

[14] N. M. Fraser and K. W. Hipel. Solving complex conflicts. *IEEE Transaction on Systems Man and Cybernetics*, 9(12):805–816, 1979.

[15] L. Gasser. Social concepts of knowledge and action: DAI foundations and open systems semantics. *Artificial Intelligence*, 47(1–3):107–138, 1991.

[16] M.R. Genesereth, M.L. Ginsberg, and J. S. Rosenschein. Cooperation without communication. In *Proc. of AAAI-86*, pages 51–57, Philadelphia, Pennsylvania, 1986.

[17] A.L. George. *Managing US-Soviet Rivalry: Problems of Crisis Prevention.* Westview Press, Boulder, 1983.

[18] A.L. George and R. Smoke. *Deterrence in American Foreign Policy.* Columbia University Press, New York, 1974.

[19] 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.

[20] M. Harris, S. Kraus, J. Wilkenfeld, and E. Blake. A decision support system for generalized negotiations. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, pages 382–387. Lawrence Erlbaum Associates, 1991.

[21] J. C. Harsanyi. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations.* Cambridge University Press, 1977.

[22] C.L Karrass. *The Negotiating Game: How to Get What You Want.* Thomas Crowell Company, NY, 1970.

[23] J. L. Kolodner and R. L. Simpson. The mediator: Analysis of an early cased-based problem solver. *Cognitive Science*, 13:507–549, 1989.

[24] S. Kraus, E. Ephrati, and D. Lehmann. Evaluation of suggestions during automated negotiations. In *Proc. of the 11th Cognitive Science Conference*, pages 299–306, 1989.

[25] S. Kraus, E. Ephrati, and D. Lehmann. Negotiation in a non-cooperative environment. *J. of Experimental and Theoretical AI*, 3(4):255–282, 1991.

[26] S. Kraus and D. Lehmann. Using interacting goal-processes for automated negotiation. In *Proc. of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 209–219, 1989.

[27] S. Kraus and D. Lehmann. Knowledge acquisition in negotiation games. In F. Geyer, editor, *The Cybernetics of Complex Systems - Self- organization, Evolution, and Social Change.* Intersystems Publications, Salinas, CA, 1991.

[28] S. Kraus and J. Wilkenfeld. Modeling a hostage crisis: Formalizing the negotiation process. Technical Report UMIACS TR 90-19 CS TR 2406, Institute for Advanced Computer Studies, University of Maryland, 1990.

[29] S. Kraus and J. Wilkenfeld. A strategic negotiations model with applications to an international crisis. Technical Report UMIACS TR 90-20 CS TR 2407, Institute for Advanced Computer Studies, University of Maryland, 1990.

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

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

[32] S. Kraus, J. Wilkenfeld, M. Harris, and B. Blake. The hostage crisis simulation. *Simulations and Games*, 23(4):398—416, 1992.

[33] T. Kreifelts and F. Von Martial. A negotiation framework for autonomous agents. In *Proc. of the Second European Workshop on Modeling Autonoumous Agents in a Multi Agent World*, pages 169–182, France, 1990.

[34] K. Kuwabara and V. Lesser. Extended protocol for multistage negotiation. In *Proc. of the Ninth Workshop on Distributed Artificiall Intelligence*, pages 129–161, Washington, 1989.

[35] B. Laasri, H. Laasri, and V. Lesser. Negotiation and its role in cooperative distributed problem solving. In *10th International Workshop on DAI*, Texas, 1990. Chapter9.

[36] R.N. Lebow. *Between Peace and War.* Johns Hopkins University Press, Baltimore, 1981.

[37] R. Leng. Structure and action in militarized disputes. In C. Hermann, C. Kegley, and J. Rosenau, editors, *New Directions in the Study of Foreign Policy.* Allen and Unwin, Winchester, MA, 1987.

[38] R. Leng. Crisis learning games. *American Political Science Review*, 82, l, 1988.

[39] R. Leng and C. Gochman. Dangerous disputes: A study of conflict behavior and war. *American Journal of Political Science*, 26(4):664–687, 1982.

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

[41] T. W. Malone, R. E. Fikes, K. R. Grant, and M. T. Howard. Enterprise: A marketlike task schedule for distributed computing environments. In B. A. Huberman, editor, *The Ecology of Computation*, pages 177–205. North Holland, 1988.

[42] S. Matwin, S. Szpakowicz, Z. Koperczak, G. E. Kersten, and W. Michalowski. Negoplan: An expert system shell for negotiation support. *IEEE Expert*, 4(4):50–62, 1989.

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

[44] J. F. Nash. Two-person cooperative games. *Econometrica*, 21:128–140, 1953.

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

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

[47] J. S. 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.

[48] A. E. Roth. *Axiomatic Models of Bargaining.* Lecture Notes in Economics and Mathematical Systems No. 170. Spring-Verlag, Berlin, 1979.

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

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

[51] 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.

[52] A. Sathi, T.E. Morton, and S.F Roth. Callisto: An intelligent project management system. *The AI Magazine*, 7(5):34–52, 1986.

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

[54] A. Shaked and J. Sutton. Involuntary unemployment as a perfect equilibrium in a bargaining model. *Econometrica*, 52(6):1351–1364, 1984.

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

[56] G.H. Snyder and P. Diesing. *Conflict Among Nations: Bargaining, Decision Making and System Structure in International Crises.* Princeton University Press, 1977.

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

[58] J.G. Stein and R. Tanter. *Rational Decision Making: Israel's Security Choices 1967.* Ohio State University Press, Columbus, 1980.

[59] 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.

[60] L. Thompson. Negotiation behavior and outcomes: Empirical evidence and theoretical issues. *Psychological Bulletin*, 108(3):515–532, 1990.

[61] K. Werkman. Knowledge-based model of negotiation using shareable perspectives. In *10th International Workshop on DAI*, Texas, 1990. Chapter 6.

[62] J. Wilkenfeld, M. Brecher, and S. Moser. *Crises in the Twentieth Century, Vol. II: Handbook of Foreign Policy Crises.* Pergamon Press, Oxford, 1988.

[63] G. Zlotkin and J. S. Rosenschein. Negotiation and task sharing among autonomous agents in cooperative domains. In *Proceedings of IJCAI-89*, pages 912–917, Detroit, Michigan, August 1989.

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