Automatica, Vol. 18, No. 4, pp. 431–441, 1982
Printed in Great Britain.

0005–1098/82/040431–11$03.00/0
Pergamon Press Ltd.

A Concept of Cooperative Equilibrium for Dynamic Games*

B. TOŁWIŃSKI†

The nonuniqueness of memory Nash equilibria in deterministic dynamic games is not a restraint for the theory but a valuable property permitting the definition of a variety of 'cheating-proof' cooperative solutions.

Key words—Game theory; differential games; dynamic programming; economics; decision theory; large-scale systems; cooperative games; bargaining models.

Abstract—The paper proposes an equilibrium solution concept for dynamic games where players can communicate with one another, but cannot make contractual agreements. In such games, unlike the static problems without contracting possibilities, the cooperation between players is possible due to the fact that the realization of negotiated agreements can be enforced by suitably-defined strategies. The definition presented combines dynamic programming, the theory of bargaining and the notion of enforceable agreements to produce a class of cooperative solutions defined in the form of memory Nash equilibria satisfying the principle of optimality along the equilibrium trajectory. The choice of a particular solution in this class depends on players' expected actions in case of disagreement, and on an adopted negotiation scheme formalized in the form of a bargaining model. Possible formulations of disagreement policies and bargaining models are discussed in some detail.

1. INTRODUCTION

In dynamic games the players' strategy spaces may include the so-called memory strategies, relating the actions of a player at a given stage of the game to the decisions made by himself and the other player at the previous stages. By declaring his memory strategy a player can determine how his future actions will depend on his opponent's prior behaviour, that is he can formulate threats or incentives aimed to induce the other player to act in some desired way. The memory strategies were first introduced in the context of 'supergames', i.e. the dynamic games consisting in repeated playing of a given static game (Aumann, 1959; Friedman, 1971, 1977). Recently, several types of memory strategies have been defined for differential games (dynamic games in the state-variable form), in the particular context of Stakelberg problems, where it is assumed that only one player, the *a priori* specified leader, is in a position to declare incentives, and that the realization of all his threats and promises is assured by a contract which is binding in the sense that once made, it literally cannot be broken (Başar and Selbuz, 1979; Tołwiński, 1979, 1980, 1981). In this paper memory strategies will be considered for discrete-time differential games without *a priori* given leaders, and with the possibility of making binding agreements ruled out. So far the most popular solution concept for the games of this type has been the pure feedback (no-memory) Nash equilibrium (Starr and Ho, 1969). Başar (1974) has observed that when in a dynamic game the players' strategy spaces include memory strategies, then as a rule the game has infinitely many Nash equilibria, from which he has concluded that a deterministic (non-hierarchical) dynamic game is not a well-posed mathematical problem (Başar, 1976). In the sequel it is shown that the deterministic formulation of dynamic game does make sense and a reasonable solution of such a game can be defined, provided that some additional characteristics of the problem are known. It is argued that the feedback Nash equilibrium is an appropriate solution concept for a strictly noncooperative game where the players are incapable not only of making binding agreements but also of communicating with one another, while memory Nash equilibria correspond to the situations where without having contracting possibilities the players can freely communicate with one another. The definition introduced in Section 3 combines the dynamic programming technique with the theory of bargaining to

*Received 16 March 1981; revised 19 October 1981; revised 23 December 1981. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by associate editor T. Başar under the direction of editor H. Kwakernaak.

†Polish Academy of Sciences, Systems Research Institute, ul. Newelska 6, 01447 Warsaw, Poland.

produce an equilibrium solution which depends on threats declared by the players and on an adopted negotiation scheme. The choice of credible threats, a particular bargaining model and the resulting equilibrium solution are discussed in Section 4. Finally, the theory is illustrated by an example in Section 5.

2. FORMULATION OF THE DYNAMIC GAME PROBLEM

Consider a deterministic dynamic system

$$x_{t+1} = f_t(x_t, u_{1t}, u_{2t}), \quad 0 \leq t < T \qquad (1)$$

where t is the integer time variable, T the time horizon (finite or infinite), $x_t \in R^n$ is the state variable, $u_{1t} \in U_{1t} \subset R^{m_1}$ and $u_{2t} \in U_{2t} \subset R^{m_2}$ are the control (decision) variables under the control of player 1 and player 2, respectively. The objective of player i, $i = 1, 2$, is the maximization of his payoff (utility) function $J_i(0, x_0; u_1, u_2)$ for a given $x_0 \in X_0$, where

$$J_i(t, x_t; u_1^t, u_2^t) = \sum_{s=t}^{T} g_{is}(x_s), \quad 0 \leq t < T \qquad (2)$$

$u_i^t = \{u_{is}\}_{s=t}^{l}$ with $l = T - 1$ if T is finite and $l = \infty$ otherwise, and $u_i = u_i^0$, $i = 1, 2$. In the sequel the following notation is used: $u_t = (u_{1t}, u_{2t})$, $u^t = (u_1^t, u_2^t)$, $u = (u_1, u_2)$, $U_i^t = \{U_{is}\}_{s=t}^{l}$, $U_t = U_{1t} \times U_{2t}$, $U^t = U_1^t \times U_2^t$ and $U = U_1 \times U_2$. Furthermore, let X_t be the reachable set at stage t, that is the set of all x_t generated by $x_0 \in X_0$ and all admissible control sequences $\{u_{i0}, u_{i1}, \ldots, u_{it-1}\}$, $i = 1, 2$. Consider dynamic games given by (1) and (2) assuming that the players have access to the closed-loop information with perfect recall, that is at every stage t the players gain information about x_t and recall $x_0, x_1, \ldots, x_{t-1}$. As a consequence, the strategy space Γ_i of player i is defined as the set of all mappings of the form

$$\gamma_i = \{\gamma_{it}\}_{t=0}^{l} \qquad (3)$$

where

$$\gamma_{it}: X_0 \times X_1 \times \ldots \times X_t \to U_{it} \qquad (4)$$

we shall also denote $\gamma_i^t = \{\gamma_{is}\}_{s=t}^{l}$, $i = 1, 2$, $\gamma_t = (\gamma_{1t}, \gamma_{2t})$, $\gamma^t = (\gamma_1^t, \gamma_2^t)$, $\gamma = (\gamma_1, \gamma_2)$ and

$$J_i(t, x_0, \ldots, x_t; \gamma^t) = \sum_{s=t}^{T} g_{is}(x_s), \quad 0 \leq t < T \qquad (5)$$

where

$$x_{s+1} = f_s(x_s; \gamma_s(x_0, x_1, \ldots, x_s)), \quad t \leq s < T \qquad (6)$$

$\gamma_i \in \Gamma_i$ is said to be a pure feedback or no-memory strategy if, for all t, γ_{it} does not depend on $x_0, x_1, \ldots, x_{t-1}$ but is a function of x_t only. Otherwise γ_i is called a memory strategy.

In the formulation of the payoff functions (2) the stage payoffs g_{it} have been assumed to be functions of the state variable only. The case of the stage payoffs having the form $h_{it}(x_t, u_t)$ can be dealt with similarly, under the additional condition that at stage $t + 1$ player i knows the value of $h_{it}(x_t, u_t)$, $i = 1, 2$. This follows from the definition of equilibrium strategies given in the next section requires player i to be able to detect any action of the other player influencing player is payoff function.

By specifying the state equation (1), the payoff functions (2) and the strategy spaces Γ_1, Γ_2 we have defined the game problem known as a two-person nonzero-sum dynamic game with the closed-loop information structure. The purpose of this paper is to discuss a solution concept for this problem.

3. EQUILIBRIUM IN DYNAMIC GAMES

The classical theory of nonzero-sum games considers two types of problems, namely the cooperative games where the players can communicate with one another and make binding agreements as how to correlate their actions, and the noncooperative games where such communication and contracting possibilities are ruled out. In the latter case the equilibrium is defined as a pair of strategies $\gamma^N = (\gamma_1^N, \gamma_2^N) \in \Gamma_1 \times \Gamma_2$ satisfying

$$J_1(\gamma^N) \geq J_1(\gamma_1, \gamma_2^N) \quad \text{for all } \gamma_1 \in \Gamma_1 \qquad (7)$$

$$J_2(\gamma^n) \geq J_2(\gamma_1^N, \gamma_2) \quad \text{for all } \gamma_2 \in \Gamma_2 \qquad (8)$$

and it is called the noncooperative Nash equilibrium. Between the extreme of purely cooperative and purely noncooperative games lies the class of problems where the players can freely communicate with one another but cannot make binding agreements (agreements between players have no legal standing or are not enforceable for other reasons). The opportunity to communicate in absence of contracting possibilities is relevant when the number of Nash equilibria is greater than one, and the players have to agree on one of them as a solution of the game. In a purely noncooperative dynamic game the lack of communication rules out memory strategies as formulation of threats is useless if they cannot be communicated to the opponent. Thus the players' strategy spaces are practically limited to include only no-memory strategies, implying that the pure feedback (no-memory) Nash equilibrium is the appropriate equilibrium concept for a strictly noncooperative dynamic game. This type of equilibrium has been given much attention in the literature, in the context of differential games, see Starr and Ho (1969) for example. In this paper, however, the main interest is dynamic games where the players can freely communicate with one another although they have no contracting possibilities. In such a case the use of memory strategies becomes essential and the definition of reasonable equilibrium less obvious. As a rule a dynamic game has infinitely many memory Nash equilibria (Başar, 1974, 1976) and the question to be answered is which one of them can be considered as a solution of the game.

First of all it is important to realize that not all Nash equilibria can be considered as candidates for equilibrium of dynamic game with contracting possibilities ruled out. In a dynamic game the players act not once but many times, facing at consecutive stages $t = 0, 1, 2, \ldots$ of the game payoff functions of the form (2). Therefore in the absence of binding agreements equilibrium strategies must be defined in the way to ensure that neither player can unilaterally improve his payoff on $[t, T]$ by changing his strategy at t, for $t = 0, 1, 2, \ldots,$ a requirement considerably stronger than conditions (7) and (8) implying only that neither player can unilaterally improve his overall payoff. Mathematically let $\bar{\gamma}_i = \{\bar{\gamma}_{it}\}_{t=0}^{l}$, $i = 1, 2$, be the equilibrium strategies and $\bar{x} = \{\bar{x}_t\}_{t=0}^{T}$ the resulting equilibrium trajectory, that is

$$\bar{x}_0 = x_0$$

$$\bar{x}_{t+1} = f_t(\bar{x}_t, \bar{\gamma}_{1t}(\bar{x}_0, \ldots, \bar{x}_t), \bar{\gamma}_{2t}(\bar{x}_0, \ldots, \bar{x}_t)). \quad (9)$$

Equilibrium strategies are required to satisfy the conditions of the form

$$J_1(t, \bar{x}_0, \ldots, \bar{x}_t;\ \bar{\gamma}_1^t, \bar{\gamma}_2^t) \geq J_1(t, \bar{x}_0, \ldots, \bar{x}_t;\ u_1^t, \bar{\gamma}_2^t)$$
for all $u_1^t \in U_1^t$ and $0 \leq t < T$.
$$(10)$$

$$J_2(t, \bar{x}_0, \ldots, \bar{x}_t;\ \bar{\gamma}_1^t, \bar{\gamma}_2^t) \geq J_2(t, \bar{x}_0, \ldots, \bar{x}_t;\ \bar{\gamma}_1^t, \bar{u}_2^t)$$
for all $u_2^t \in U_2^t$ and $0 \leq t < T$.
$$(11)$$

In other words, the set of candidates for equilibrium solution is restricted to include only Nash strategies satisfying the principle of optimality along the equilibrium trajectory. Now the problem is to select, from this still an infinite set, an element which can be considered as an equilibrium solution of the game. The proposed approach to this problem combines the dynamic programming with the theory of bargaining and can be summarized as follows: First the players define their disagreement (threat) policies which they are supposed to use if they cannot agree on another solution, or if after reaching an agreement one of them breaks it in the course of the game. Note that the disagreement policies play a dual role, on the one hand they define reference points for the negotiation of agreements at consecutive stages, and on the other hand they define threats discouraging the players from breaking the agreements in the course of the game.* Given threat policies, equilibrium decisions (agreements) are determined stage by stage backward in time, that is according to the dynamic programming scheme. At the last stage of the game (in the case when T is finite and the game has the last stage) equilibrium decisions are defined as non-cooperative Nash strategies of the corresponding static game. At all other stages they are selected by means of a given bargaining model from the set of enforceable agreements, where an agreement at stage t is said to be enforceable if neither player can gain anything from breaking it, assuming that such an action will be followed by the switch to disagreement policies from stage $t+1$ onwards. The resulting equilibrium strategies have the form

$$\gamma_{i0}(x_0) = \bar{u}_{i0}$$

$$\gamma_{it}(x_0, \ldots, x_t) = \begin{cases} \bar{u}_{it} \text{ if } x_1 = \bar{x}_1, x_2 = \bar{x}_2, \ldots, x_t = \bar{x}_t \\ d_{it}(x_t) \text{ otherwise} \end{cases} \quad (12)$$

$$0 \leq t < T; \quad i = 1, 2$$

where $\bar{u}_{1t}$, $\bar{u}_{2t}$ denote equilibrium decisions at stage t, $\{x_0, \bar{x}_1, \ldots, \bar{x}_T\}$ is the corresponding equilibrium trajectory obtained by substituting $\bar{u}_{1t}, \bar{u}_{2t}$ for u_{1t}, u_{2t} in the state equation, and $d_{it}(x_t)$ is the disagreement action of player i at stage t for $x_t \in X_t$, $i = 1, 2$, $0 \leq t < T$. By declaring a strategy of the form (12) player i commits himself to the possible realization of

*This property is not crucial for the proposed equilibrium concept which can also be defined in a more general way, allowing for the negotiation threats used to reach an agreement to be different from the retaliation threats used to prevent the violation of the agreement which has already been negotiated.

the treat d_{it} which need not to be optimal in terms of his payoff function, that is after $x_t \neq \bar{x}_t$ player i may be better off by not using the disagreement policy. The rationale for such a commitment is to achieve favorable cooperative controls $(\bar{u}_{1t}, \bar{u}_{2t})$ which depend on the disagreement policies through the bargaining model and the definition of the set of enforceable agreements. Some possible formulations of d_{it} will be discussed in the next section.

Assuming that disagreement actions $d_{it}(x)$ for $x_t \in X_t$, $0 \leq t < T$, $i = 1, 2$ are given and a bargaining model F is specified, where F is defined as a mapping assigning to a static cooperative game G and a disagreement point d the bargaining solution $F(G, d)$, the definition of equilibrium solution of a dynamic game can be formally introduced in the following way: Let $e_{1t}(x_t)$ and $e_{2t}(x_t)$ denote equilibrium decisions (agreements) of player 1 and player 2, respectively, for $x_t \in X_t$ and $t = 0, 1, 2, \ldots$. Denote also

$$e_i^t(x_t) = \{e_{is}(x_s)\}_{s=t}^{l}, \quad i = 1, 2 \tag{13}$$

$$d_i^t(x_t) = \{d_{is}(x_s)\}_{s=t}^{l}, \quad i = 1, 2 \tag{14}$$

with

$$x_{s+1} = f_s(x_s, e_{1s}(x_s), e_{2s}(x_s))$$
$$x_{s+1} = f_s(x_s, d_{1s}(x_s) d_{2s}(x_s))$$
$$\text{for } s = t, t+1, \ldots \tag{15}$$

respectively

$$D_{1t}(x_t) = \sup_{u_1^t} J_1(t, x_t;\ u_1^t, d_2^t(x_t))$$

$$D_{2t}(x_t) = \sup_{u_2^t} J_2(t, x_t;\ d_1^t(x_t), u_2^t)$$
$$\text{for } t = 0, 1, \ldots, T-1 \tag{16}$$

$$H_{1t}(x_t, u_{2t}) = \sup_{u_{1t}} [g_{1t}(x_t) + D_{1t+1}(f_t(x_t, u_{1t}, u_{2t}))]$$

$$H_{2t}(x_t, u_{1t}) = \sup_{u_{2t}} [g_{2t}(x_t) + D_{2t+1}(f_t(x_t, u_{1t}, u_{2t}))]$$
$$\text{for } t = 0, 1, \ldots, T-2 \tag{17}$$

$$e_t = (e_{1t}, e_{2t}), \quad e^t = (e_1^t, e_2^t), \quad d_t = (d_{1t}, d_{2t}), \quad d^t = (d_1^t, d_2^t) \tag{18}$$

$$\bar{J}_{it}(x_t, u_t) = g_{it}(x_t) + J_i(t+1, x_{t+1}, e^{t+1}(x_{t+1})) \tag{19}$$

where $x_{t+1} = f_t(x_t, u_t)$.

We can easily interpret the expression $D_{it}(x_t)$ given in (16) as the maximum payoff that player i can secure the stages t to T given that the system is in state x_t at t, and that player j will play according to his threat policy d_j, $j \neq i$. The expression $H_{it}(x_t, u_{jt})$ given in (18) is the maximum payoff that player i may expect for the stages t to T if the system is in state x_t at t, and the opponent is bound to play u_{jt} at t but will use his threat policy from stage $t+1$ onward. This expression plays a fundamental role in the definition of the set of enforceable agreements which for a given x_t at t and given equilibrium decisions from stage $t+1$ onward is defined as

$$A_t(x_t, d^{t+1}) = \{u_t \in U_t / \bar{J}_{it}(x_t, u_t) \geq H_{it}(x_t, u_{jt}), \quad i = 1, 2, j \neq i\} \tag{20}$$

for $t = 0, 1, \ldots, T-2$ if T is finite and $t = 0, 1, 2, \ldots$ otherwise. In other words, if $\bar{u}_t = (\bar{u}_{1t}, \bar{u}_{2t})$ is an agreement reached at stage t, and $\bar{u}_t \in A_t(x_t, d^{t+1})$, then neither player can benefit from cheating provided the compliance with the agreement is followed by the realization of $e^{t+1}(x_{t+1})$, and noncompliance by the execution of the opponent's threat d_j^{t+1}. Given $0 \leq t \leq T-2$ and $x_t \in X_t$, and assuming that the equilibrium strategies will be applied in the future the players face a static cooperative game $G_t(x_t, d^{t+1})$ defined by $A_t(x_t, d^{t+1})$ as the set of admissible solutions and $\bar{J}_{1t}(x_t u_t)$, $\bar{J}_{2t}(x_t; u_t)$ as payoff functions, with $u_t \in A_t(x_t, d^{t+1})$. The last stage of the game is different because decisions made at this stage cannot be influenced by any threats to be realized in the future, so if T is finite then at $t = T-1$ the players face a noncooperative static game $G_{T-1}(x_{T-1})$ defined by decision sets U_{1T-1}, U_{2T-1} and payoff functions

$$\bar{J}_{1T-1}(x_{T-1};\ u_{1T-1}, u_{2T-1})$$
$$\bar{J}_{2T-1}(x_{T-1};\ u_{1T-1}, u_{2T-1}) \text{ with } u_{iT-1} \in U_{iT-1}, \quad i = 1, 2.$$

Now, the assumption that the bargaining model F is well defined on pairs (G_t, d^t) leads to the following definition:

Definition 1

(1) Let for any $0 \leq t < T$, with the possible exception of $t = T-1$ in the case when T is finite, and any $x_t \in X_t$

$$e_t(x_t) = F(G_t(x_t, d^{t+1}), d^t) \tag{21}$$

(2) If T is finite, then let for any $x_{T-1} \in X_{T-1}$, $e_{T-1}(x_{T-1})$ be the noncooperative Nash equilibrium of $G_{T-1}(x_{T-1})$.

Furthermore let

$$\bar{x}_{t+1} = f_t(\bar{x}_t, e_t(\bar{x}_t)) \quad 0 \leq t < T, \quad \bar{x}_0 = x_0 \qquad (22)$$

and

$$\bar{u}_t = (\bar{u}_{1t}, \bar{u}_{2t}) = (e_{1t}(\bar{x}_t), e_{2t}(\bar{x}_t)), \quad 0 \leq t < T. \qquad (23)$$

Then the strategy pair (γ_1, γ_2) with γ_{it} defined by (12) is said to be the equilibrium solution of the closed-loop dynamic game formulated in the previous section, under the condition that the players have no contracting possibilities but can freely communicate with one another. □

The definition of $\bar{u}_t$ as enforceable agreements at $t = 0, 1, \ldots T-2$ and as the Nash equilibrium at $t = T-1$ implies that the equilibrium strategies (12) satisfy the conditions (10), (11), i.e. the principle of optimality along the equilibrium trajectory. Assuming F to produce bargaining solutions which are Pareto optimal implies $\bar{u}_t$ to be Pareto optimal for the local games $G_t(x_t, d^{t+1})$, but, of course, does not imply $\bar{u}$ to be globally Pareto optimal for the dynamic game as a whole. Finally it should be noted that the equilibrium solution produced by Definition 1 depends in a crucial way on the choice of disagreement policies and bargaining model. In the next section possible formulations of d_{it} and F will be discussed.

4. DISAGREEMENT POLICIES AND BARGAINING MODELS

The equilibrium solution introduced in the previous section assumes the existence of threat policies which are credible in the sense that each player believes that if he fails to cooperate then the other player will actually realize his threat. The choice of threats has a crucial impact on the players' payoffs under the equilibrium solution, because the equilibrium agreements $\bar{u}_t$ resulting from Definition 1 depend on d^t through both the bargaining model F and the set A_t. The definition of optimal or credible bargaining threats is a well-known problem of the game theory, and has been discussed in the context of static problems by Nash (1953), Luce and Raiffa (1957) and Rosenthal (1976) among others. In the sequel two formulations of threats for the dynamic game of Section 2 will be considered, namely as pure feedback Nash strategies, and as minimax strategies of a zero-sum game. The first approach is based on the hypothesis that if the players fail to cooperate then they will switch to the purely noncooperative equilibrium, implying d_1^t, d_2^t to be defined as no-memory Nash strategies. In such a case an equilibrium solution resulting from Definition 1 has an attractive property of fulfilling an even stronger version of conditions (10), (11), namely

$$J_1(t, x_0, \ldots, x_t; \gamma_1^t(x_t), \gamma_2^t(x_t)) \geq J_1(t, x_0, \ldots, x_t; u_1^t, \gamma_2^t(x_t)) \qquad (24)$$

for all $u_1^t \in U_1^t$, any $x_t \in X_t$, and $0 \leq t < T$

$$J_2(t, x_0, \ldots, x_t; \gamma_1^t(x_t), \gamma_2^t(x_t)) \geq J_2(t, x_0, \ldots, x_t; \gamma_1^t(x_t), u_2^t) \qquad (25)$$

for all $u_2^t \in U_2^t$, any $x_t \in X_t$, and $0 \leq t < T$.

In other words, the strategy γ_i^t is optimal for the player, provided the other one applies γ_j^t, $j \neq i$, independently of whether the cooperation succeeded or not in the past. Clearly the threat policies defined as pure feedback Nash strategies are quite credible. On the other hand Definition 1 with such threat policies can produce a non-trival equilibrium solution only for games with the infinite time horizon (Friedman, 1977), because if T is finite then the sets of enforceable agreements reduce to include only noncooperative Nash solutions and the resulting equilibrium solution of the game is simply the feedback Nash equilibrium. Another problem is that Nash strategies may not be satisfactory as the reference points in the bargaining, especially when the comparisons between the players' utility functions are acceptable (Luce and Raiffa, 1957). In such a case it makes sense to consider the zero-sum game defined by the payoff function of the form

$$J(0, x_0; u) = J_2(0, x_0; u) - J_1(0, x_0; u). \qquad (26)$$

and the disagreement policies d_1^t, d_2^t can be formulated as the minimax strategies, provided of course that such strategies exist.* In other words one has

$$J(t, x_t; d_1^t(x_t), d_2^t(x_t)) = \inf_{u_1^t \in U_1^t} \sup_{u_2^t \in U_2^t} J(t, x_t; u_1^t, u_2^t)$$

$$= \sup_{u_2^t \in U_2^t} \inf_{u_1^t \in U_1^t} J(t, x_t; u_1^t, u_2^t) \stackrel{\Delta}{=} M_t(x_t) \qquad (27)$$

where

$$J(t, x_t; u_1^t, u_2^t) = J_2(t, x_t; u_1^t, u_2^t) - J_1(t, x_t; u_1^t, u_2^t). \qquad (28)$$

*For the justification of this definition see the comment made in this section after the discussion of the Raiffa's bargaining model.

Note that for $t = T-1, T-2, \ldots, 1$, d_{1t} and d_{2t} can be obtained by means of the dynamic programming rule

$$M_t(x_t) = g_{2t}(x_t) - g_{1t}(x_t) + M_{t+1}[f_t(x_t, d_{1t}(x_t), d_{2t}(x_t))]$$
$$= g_{2t}(x_t) - g_{1t}(x_t) + \inf_{u_{1t} \in U_{1t}} \sup_{u_{2t} \in U_{2t}} M_{t+1}(f_t(x_t, u_{1t}, u_{2t})). \quad (29)$$

An equilibrium solution with the minimax strategies of (26) as the threat policies assumes that the failure of cooperation makes the players adopt the strictly antagonistic attitude, which is a change from their original objectives. It seems evidence of such behaviour can be found if one takes a close look at some real-life games, as for example relations between the great powers.

The next problem to be considered in connexion with Definition 1 is the choice of a bargaining model F. Traditionally (Roth, 1979) a bargaining game is defined as a pair (S, m) where S is a subset of R^2 representing the feasible payoff pairs to the players, and m is an element of S corresponding to the disagreement outcome. Let B be a class of bargaining games (S, m). A bargaining model for the class B is defined as a function F': $B \to R^2$ such that $F'(S, m)$ is an element of S for any (S, m) in B (Roth, 1979). The relations between bargaining games (G, d) and (S, m), and bargaining models F and F' are straightforward. If G is described by payoff functions J_1, J_2 and decision sets U_1, U_2, while $d = (d_1, d_2) \in U_1 \times U_2$ represents the disagreement actions then

$$S = \{y = (y_1, y_2) \in R^2 / \exists u \in U_1 \times U_2, \ y_1 = J_1(u) \text{ and } y_2 = J_2(u)\} \quad (30)$$

$$m = (J_1(d), J_2(d)). \quad (31)$$

Furthermore, if for some $u \in U_1 \times U_2$ we have $(J_1(u), J_2(u)) = F'(S, m)$ then $u = F(G, d)$.

Note that a well-defined bargaining model should produce for any game in B a solution which is unique in terms of F' but not necessarily unique in terms of F.

The problem of bargaining has been extensively discussed in the game theory literature (Nash, 1950; Raiffa, 1953; Luce and Raiffa, 1957; Roth, 1979). The best known is the Nash bargaining model $F' = N'$ defined by $N'(S, m) = y$ such that $y \geqslant m$ and

$$(y_1 - m_1)(y_2 - m_2) > (z_1 - m_1)(z_2 - m_2)$$

for all z in S such that $z \geqslant m$ and $z \neq y$. N' possesses four important properties, namely independence of equivalent utility representations, symmetry, independence of irrelevant alternatives and Pareto optimality, and is well defined on the class of games with S being compact, convex and including at least one point $s \in S$ such that $m < s$. Another bargaining model has been proposed by Raiffa (1953) for cases where interpersonal utility comparisons make sense. Let $P(S)$ denote the Pareto set, that is

$$P(S) = \{y \in S | z \geqslant y \text{ and } z \in S \text{ implies } z = y\}. \quad (32)$$

The bargaining model $F' = R'$ is defined by $R'(S, m) = y$ such that

$$\min\{y_1 - m_1, y_2 - m_2\} > \min\{z_1 - m_1, z_2 - m_2\} \quad (33)$$

for all $z \in P(S)$ such that $z \neq y$ (Roth, 1979).

R' always chooses the Pareto optimal point which comes closest to giving the players equal gains, and is well defined if $P(S)$ is compact and contractible (i.e. it contains no 'holes'). If there is a point y in $P(S)$ such that $y_1 - m_1 = y_2 - m_2$, then $R'(S, m) = y$, that is the selected solution gives both players equal gains, whenever there is Pareto optimal outcome with this property. If the straight line $y_1 - y_2 = m_1 - m_2$ does not have a common point with $P(S)$, then $P(S)$ is contained either in the set $\{y | y_1 - m_1 > y_2 - m_2\}$ or in the set $\{y | y_1 - m_1 < y_2 - m_2\}$. In the first case $R(S, m) = y$ is the point at which y_2 is maximized, in the second case it is the point which maximizes y_1.

The transition from the formulation of the models in terms of F' to the formulation in terms of F required by Definition 1 is straightforward in case of N', but needs some comments in case of R'. Let

$$J(u) = J_2(u) - J_1(u), \quad M = m_2 - m_1 \quad (34)$$

and define

$$W_1 = \{u \in U_1 \times U_2 | J(u) \geqslant M\} \quad (35)$$

$$W_2 = \{u \in U_1 \times U_2 | J(u) \leqslant M\} \quad (36)$$

$$Y_i = \{v \in W_i | J_i(v) \geqslant J_i(u) \text{ for all } u \in W_i\}, \quad i = 1, 2 \quad (37)$$

$$v^1 = \arg\max_{u \in Y_1} J_2(u) \quad (38)$$

$$v^2 = \arg\max_{u \in Y_2} J_1(u). \quad (39)$$

The interpretation of the set W_i is that it is the set of agreements which, if proposed by player i, should be accepted by the other player, because in the case of non-acceptance followed by the realization of disagreement payoffs the relative loss of the latter would be greater or equal to the relative loss of player i. So, v^i is by definition an optimal choice of player i in the set of agreements that the other player is expected to accept, having the additional property that when such a choice is not unique v^i is selected to maximize the payoff of player $j, j \neq i$. Now it can be easily seen that the bargaining model R can be defined by $R(G, d) = v^i$ where $i = 1$ if

$$J_k(v^1) \geq J_k(v^2), \quad k = 1, 2 \tag{40}$$

and $i = 2$ if

$$J_k(v^2) \leq J_k(v^1), \quad k = 1, 2. \tag{41}$$

Note that if the line of equal gains $y_2 - y_1 = M$ intersects $P(S)$ then $J_k(v^1) = J_k(v^2)$ for $k = 1, 2$, if $P(S)$ lies above this line then condition (40) is satisfied, and if $P(S)$ lies below it then one has (41). The model R has a particularly interesting interpretation when the disagreement point d is defined as a minimax solution of J, that is

$$M = J(d) = \min_{u_1 \in U_1} \max_{u_2 \in U_2} J(u_1, u_2). \tag{42}$$

In such a case the set

$$W_i^0 = W_i/\{u \in U_1 \times U_2 | J(u) = M\} \tag{43}$$

can be considered as the set of agreements with respect to which player i can formulate the threat d_i, which is credible in the sense that for any possible counteraction of the other player the relative loss in the payoff of the latter will be greater than the relative loss in payoff of player i. Neither player can formulate credible threats with respect to points in the set $\{u \in U_1 \times U_2 | J(u) = M\}$, so the bargaining solution is selected as an agreement maximizing the players' payoffs on this set, unless both players can benefit from an agreement on a point $u \in W_i^0$, for $i = 1$ or $i = 2$, and then the solution is chosen as a point that maximizes J_i on W_i^0.

A given combination of threat policies and of a bargaining model can produce the equilibrium solution of a dynamic game, provided the required policies exist and the bargaining model is well defined on the family of games $G_t(x_t; d^{t+1})$, where $t = 0, 1, \ldots, T-2$ and $x_t \in X_t$. As for the threat policies considered in this section, there are very few results of practical value concerning the existence or computation of no-memory Nash equilibria, but on the other hand, the minimax problem is more tractable, for instance the existence of minimax strategies is ensured if the sets U_{it}, $i = 1, 2$, $0 \leq t < T$ are compact and convex, and the function J is continuous, convex with respect to u_1 and concave with respect to u_2 (Owen, 1968). Some insight into the problem of existence and computation of equilibrium agreements in the sense of model R can be gained by formulating optimization problems with the property that for an important class of games their solutions coincide with the equilibrium agreements generated by Definition 1 with R as a bargaining model. Assume that the threat actions $d_{it}(x_t)$ for $x_t \in X_t$, $i = 1, 2$; $t = 1, 2, \ldots$ are given, and let $J(t, x_t; u^t)$, $M_t(x_t)$, $H_{1t}(x_t, u_{2t})$ and $H_{2t}(x_t, u_{1t})$ be defined by (28), (27) and (17), respectively. The following optimization problems are defined:

(a) To maximize $J_1(s, x_s; u^s)$ subject to the state equation

$$x_{t+1} = f_t(x_t, u_t), \quad t = s, s+1, \ldots, \tag{44}$$

with x_s given and subject to the following constraints

$$J(t, x_t; u^t) \geq M_t(x_t), \quad t = s, s+1, \ldots, T-2 \tag{45}$$

$$J_1(t, x_t; u^t) \geq H_{1t}(x_t, u_{2t}), \quad t = s, s+1, \ldots, T-2 \tag{46}$$

$$J_2(t, x_t; u^t) \geq H_{2t}(x_t, u_{1t}), \quad t = s, s+1, \ldots, T-2 \tag{47}$$

$$u_{T-1} = e_{T-1}(x_{T-1}) \quad \text{(only if } T \text{ is finite)} \tag{48}$$

[Recall that $e_{T-1}(x_{T-1})$ is the noncooperative Nash equilibrium of $G_{T-1}(x_{T-1})$];

(b) to maximize $J_2(s, x_s; u^s)$ subject to (44), (46)–(48), and

$$J(t, x_t; u^t) \leq M_t(x_t), \quad t = s, s+1, \ldots, T-2. \tag{49}$$

Note that in the case of finite T, the set of sufficient conditions for the existence of solutions to (a) and (b) is compactness of U_{it}, and continuity of f_t, g_{it}, d_t and e_{T-1}. Let, for given s and x_s, $a^s(x_s)$ and $b^s(x_s)$ be optimal solutions to the optimization problems (a) and (b), respectively. Denote $u^a = a^0(x_0)$, $u^b = b^0(x_0)$, and let x^a and x^b be the corresponding trajectories. The following proposition is a direct consequence of the definition of the bargaining model R.

Proposition 1

(1) If $x^a = x^b = \bar{x}$, then

$$J(t, \bar{x}_t;\, u^{at}) = J(t, \bar{x}_t;\, u^{bt}) = M_t(\bar{x}_t),\quad t = 0, 1, \ldots, T-2 \tag{50}$$

and u^a, as well as u^b, define the sequence of equilibrium agreements in the sense of model R.

(2) If the optimal trajectory x^a is unique and

$$J_k(t, x_t^a;\, u^{at}) \geqslant J_k(t, x_t^a;\, b^t(x_t^a)) \tag{51}$$

for $t = 0, 1, \ldots, T-2$ and $k = 1, 2$ then u^a defines the sequence of equilibrium agreements in the sense of model R.

(3) If the optimal trajectory x^b is unique and

$$J_k(t, x_t^b;\, u^{bt}) \geqslant J_k(t, x_t^b;\, a^t(x_t^b)) \tag{52}$$

for $t = 0, 1, \ldots, T-2$ and $k = 1, 2$ then u^b defines the sequence of equilibrium agreements in the sense of model R.

Proposition 1 provides a method for computing the equilibrium agreements in the case when the Pareto sets intersect the lines of equal gains at all stages of the game, and in the case when the Pareto sets lie always at the same side of such lines.

In this section possible realizations of the general equilibrium concept introduced by Definition 1 have been discussed. It should be noted that one such realization has been considered by Friedman (1971, 1977) within the context of the so-called supergames, which can be formulated as the infinite horizon dynamic games with the state equation of the form

$$x_{t+1} = f(u_{1t}, u_{2t}). \tag{53}$$

implying that the payoffs at stage t do not depend on past decisions. For this class of problems Friedman introduced an equilibrium concept (the 'balanced temptation equilibrium') which is also the equilibrium solution in the sense of Definition 1, with feedback Nash strategies as the disagreement policies and a specific bargaining scheme utilizing the particular form of payoff functions. As Friedman (1977) points out the balanced temptation equilibrium does not seem to make sense within a more general dynamic game context.

5. EXAMPLE

As an illustration of the proposed solution concept consider equilibrium solutions of a simple game with infinite time horizon, linear state equation of the form

$$\begin{aligned} x_{1t+1} &= ax_{1t} + u_{1t}, \quad u_{1t} \geqslant 0 \\ x_{2t+1} &= ax_{2t} + u_{2t}, \quad u_{2t} \geqslant 0 \end{aligned} \tag{54}$$

and quadratic payoff functions defined as

$$\begin{aligned} J_1(t, x_t;\, u^t) &= \sum_{s=t}^{\infty} (1+\lambda)^{-s}[(x_{1s} - x_{2s}) - (u_{1s} + u_{1s}^2)] \\ J_2(t, x;\, u^t) &= \sum_{s=t} (1+\lambda)^{-s}[(x_{2s} - x_{1s}) \\ &\quad - (u_{2s} + (1/c)u_{2s}^2)] \end{aligned} \tag{55}$$

where x_{1t}, x_{2t}, u_{1t} and u_{2t} are scalar variables, and a, c and λ are real parameters satisfying $0 < a < 1$, $0 < c \leqslant 1$ and $0 < \lambda < a$. This game may be viewed as a very simple model of the arms race between two countries, where x_{it} denotes the war potential of country i, and each country is assumed to maximize its superiority (or minimize the inferiority) minus the cost of armaments. If $c < 1$ then the unit cost of increasing the war potential is greater for country 2 than for country 1, putting the latter in a privileged position. Depreciation and discounting parameters are denoted by a and λ, respectively. The fact that the payoff functions are not of the form (2) is without consequence here, because knowing x_t, x_{t+1} and u_{it} player i can determine the decision made by the other player at stage t, and thus verify whether the latter has or has not complied with the equilibrium agreement. We shall calculate the equilibrium solutions of the game with feedback Nash and minimax strategies as the threat policies, and R as the bargaining model.

Let $\mu = 1/(1+\lambda-a)$. It is easy to see that the payoff functions can be represented in the form

$$\begin{aligned} J_1(t, x_t;\, u^t) &= \mu(1+\lambda)^{-t+1}(x_{1t} - x_{2t}) \\ &\quad + \sum_{s=t} (1+\lambda)^{-s}[\mu(u_{1s} - u_{2s}) - (u_{1s} + u_{1s}^2)] \end{aligned} \tag{56}$$

$$\begin{aligned} J_2(t, x_t;\, u^t) &= \mu(1+\lambda)^{-t+1}(x_{2t} - x_{1t}) \\ &\quad + \sum_{s=t} (1+\lambda)^{-s}[\mu(u_{2s} - u_{1s}) - (u_{2s} + (1/c)u_{2s}^2)] \end{aligned} \tag{57}$$

$$\begin{aligned} J(t, x_t;\, u^t) &= 2\mu(1+\lambda)^{-t+1}(x_{2t} - x_{1t}) \\ &\quad + \sum_{s=t} (1+\lambda)^{-s}[2\mu(u_{2s} - u_{1s}) + (u_{1s} + u_{1s}^2) \\ &\quad - (u_{2s} + (1/c)u_{2s}^2)]. \end{aligned} \tag{58}$$

The Nash and minimax strategies (denoted by $d_{it}^{(n)}(x_t)$ and $d_{it}^{(m)}(x_t)$, respectively) can be obtained from the equations

$$(\partial/\partial u_{1t})J_1(t, x_t; u^t) = 0$$
$$(\partial/\partial u_{2t})J_2(t, x_t; u^t) = 0 \tag{59}$$

and

$$(\partial/\partial u_{1t})J(t, x_t; u^t) = 0$$
$$(\partial/\partial u_{2t})J(t, x_t; u^t) = 0 \tag{60}$$

respectively, resulting in

$$d_{1t}^{(n)}(x_t) = d_1^{(n)} = (\mu - 1)/2$$
$$d_{2t}^{(n)}(x_t) = d_2^{(n)} = c(\mu - 1)/2 \tag{61}$$

and

$$d_{1t}^{(m)}(x_t) = d_1^{(m)} = (2\mu - 1)/2$$
$$d_{2t}^{(m)}(x_t) = d_2^{(m)} = c(2\mu - 1)/2 \tag{62}$$

for any $x_t \in X_t$ and $t = 0, 1, 2, \ldots$. Furthermore one has

$$M_t^{(n)} = J(t, x_t; d^{(n)t}) = -(1/4\lambda(1+\lambda)^{t-1})(\mu - 1)$$
$$(3\mu - 1)(1 - c) + 2\mu(1+\lambda)^{-t+1}(x_{2t} - x_{1t}) \tag{63}$$

and

$$M_t^{(m)} = J(t, x_t; d^{(m)t}) = -(1/4\lambda(1+\lambda)^{t-1})(2\mu - 1)^2$$
$$(1 - c) + 2\mu(1+\lambda)^{-t+1}(x_{2t} - x_{1t}). \tag{64}$$

Consider the maximization of $J_1(0, x_0; u)$ subject to the constraint

$$J(t, x_t; u^t) \geq M_t^{(n)}, \quad t = 0, 1, 2, \ldots \tag{65}$$

and the maximization of $J_2(0, x_0; u)$ subject to

$$J(t, x_t; u^t) \leq M_t^{(n)}, \quad t = 0, 1, 2, \ldots. \tag{66}$$

It is easy to see that optimal solutions of both problems are identical and given by

$$\bar{u}_{1t}^{(n)} = \bar{u}_1^{(n)} = (1/2)[2\mu - 1$$
$$- \sqrt{((\mu - 1)(1 - 3\mu)(1 - c) + (2\mu - 1)^2)}]$$
$$\bar{u}_{2t}^{(n)} = \bar{u}_2^{(n)} = 0, \quad t = 0, 1, 2, \ldots \tag{67}$$

where the control sequence consisting of controls $\bar{u}^{(n)} = (\bar{u}_1^{(n)}, \bar{u}_2^{(n)})$ is on the equal gains line, that is

$$J(t, x_t; \bar{u}^{(n)t}) = M_t^{(n)}, \quad t = 0, 1, 2. \tag{68}$$

Moreover, it can be verified that the solution (67) is enforceable, that is it satisfies the constraints (46) and (47). Thus, by Proposition 1 $\bar{u}^{(n)}$ defines the sequence of equilibrium agreements in the sense of bargaining model R. Observe that in this case the sequence of equilibrium agreements is in the class of globally Pareto optimal solutions. The equilibrium strategies in the sense of Definition 1 was given by

$$\gamma_{i0}^{(n)}(x_0) = \bar{u}_i^{(n)}$$
$$\gamma_{it}^{(n)}(x_0, \ldots, x_t) = \begin{cases} \bar{u}_i^{(n)} & \text{if } x_1 = \bar{x}_1^{(n)}, \ldots, x_t = \bar{x}_t^{(n)} \\ d_i^{(n)} & \text{otherwise} \end{cases} \tag{69}$$
$$t = 1, 2, 3, \ldots, i = 1, 2.$$

The case of $d_i^{(m)}$ as the threat policies can be dealt with analogously. From the maximization of $J_1(0, x_0; u)$ subject to

$$J(t, x_t; u^t) \geq M_t^{(m)} \tag{70}$$

this leads to

$$\bar{u}_{1t}^{(m)} = \bar{u}_1^{(m)} = d_1^{(n)} = (\mu - 1)/2, \quad \bar{u}_{2t}^{(m)} = \bar{u}_2^{(m)} = 0 \tag{71}$$

if

$$0 < c < \mu^2/2\mu - 1)^2 = 1/(1 + a - \lambda)^2 \tag{72}$$

and

$$\bar{u}_{1t}^{(m)} = \bar{u}_1^{(m)} = (2\mu - 1)(1 - \sqrt{(c)})/2$$
$$\bar{u}_{2t}^{(m)} = \bar{u}_2^{(m)} = 0 \tag{73}$$

if

$$1/(1 + a\lambda)^2 \leq c \leq 1. \tag{74}$$

In the first case

$$J(t, x_t; u^{(m)t}) > M_t^{(m)}, \quad t = 0, 1, 2, \ldots \tag{75}$$

while in the second case the control sequence defined by (73) is on the equal gains lines. Both (71) and (73) define Pareto optimal solutions. It can be shown that the solution (71) is enforceable under (72) for arbitrary a and λ, while (73) is enforceable for all c satisfying (74) provided that

$$a < (1 + \lambda^2)/(1 + \lambda). \tag{76}$$

Using Proposition 1 it is easy to verify that (71) and (73) define the sequence of equilibrium agreements in the sense of model R, for the parameters a, c, λ satisfying (72), and (74) and (76), respectively. The resulting equilibrium strategies, namely

$$\gamma_{i0}^{(m)} = \bar{u}_i^{(m)}$$

$$\gamma_{it}^{(m)}(x_0, \ldots, x^t) = \begin{cases} \bar{u}_i^{(m)} & \text{if } x_1 = \bar{x}_1^{(m)}, \ldots, x_t = \bar{x}_t^{(m)} \\ d_i^{(m)} & \text{otherwise.} \end{cases} \quad (77)$$

improve the payoff of player 1 at the expense of player 2, when compared with the solution (69). In particular, when c is sufficiently small a better bargaining position enablees player 1 to achieve the team optimal solution (71).

In the context of the arms race problem the solutions (69) and (77) can be interpreted in the following way. When the competing countries do not communicate with one another, then they can be expected to act according to the no-memory Nash equilibrium $(d_1^{(n)}, d_2^{(n)})$ defined by (61). This situation changes when the communication is established, and the countries can enter upon negotiations to limit the arms race and achieve equilibrium at lower level of military spending. The arms limitation treaty $\bar{u}^{(n)}$ takes $(d_1^{(n)}, d_2^{(n)})$ as the reference point for bargaining, and reduces the military spending of both countries, granting them equal gains measured in terms of their utility functions. However, country 1 having an edge on country 2 by producing its armaments more efficiently, can achieve an even better treaty $\bar{u}^{(m)}$ by declaring the threat $d_1^{(m)}$ which is tougher, that is implies a quicker pace of arms race in the case of unsuccessful negotiations, than the pure feedback Nash strategy $d_1^{(n)}$. If country 2 has reasons to believe that country 1 is really committed to realize $d_1^{(m)}$, it has little choice but accept $\bar{u}^{(m)}$. Table 1 shows the numerical values of $d^{(n)}$, $d^{(m)}$, $\bar{u}^{(n)}$, $\bar{u}^{(m)}$ and the corresponding payoffs for $a = 0.9$, $\lambda = 0.1$, and c ranging from 0.1 to 1. Note, that for $c = 0.1$ and 0.3 the payoffs of player 2 corresponding to $u^{(m)}$ are smaller than his payoffs corresponding to $d^{(n)}$, that means for a 'weak' player the possibility of communication can be harmful.

TABLE 1. EQUILIBRIUM SOLUTIONS OF THE ARMS RACE MODEL FOR $a = 0.9$, $\lambda = 0.1$, $x_{10} = x_{20}$, AND VARIOUS VALUES OF c

	u_1	u_2	$J_1(u_1, u_2)$	$J_2(u_1, u_2)$
$c = 0.1$				
$d^{(n)}$	2	0.2	33	−105.6
$d^{(m)}$	4.5	0.45	−49.5	−249.975
$\bar{u}^{(n)}$	1.734137	0	43.222485	−95.377518
$\bar{u}^{(m)}$	2	0	44	−110
$c = 0.3$				
$d^{(n)}$	2	0.6	11	−96.8
$d^{(m)}$	4.5	1.35	−99	−254.925
$\bar{u}^{(n)}$	1.267354	0	38.09553	−69.70447
$\bar{u}^{(m)}$	2	0	44	−110
$c = 0.5$				
$d^{(n)}$	2	1	−11	−88
$d^{(m)}$	4.5	2.25	−148.5	−259.875
$\bar{u}^{(n)}$	0.859945	0	29.70302	−47.29698
$\bar{u}^{(m)}$	1.318019	0	38.883933	−72.491061
$c = 0.7$				
$d^{(n)}$	2	1.4	−33	−79.2
$d^{(m)}$	4.5	3.15	−198	−264.825
$\bar{u}^{(n)}$	0.493755	0	19.043481	−27.156519
$\bar{u}^{(m)}$	0.735023	0	26.398355	−40.426645
$c = 0.9$				
$d^{(n)}$	2	1.8	−55	−70.4
$d^{(m)}$	4.5	4.05	−247.5	−269.775
$\bar{u}^{(n)}$	0.158341	0	6.691226	−8.708774
$\bar{u}^{(m)}$	0.230925	0	9.574115	−12.700885
$c = 1.0$				
$d^{(n)}$	2	2	−66	−66
$d^{(m)}$	4.5	4.5	−272.25	−272.25
$\bar{u}^{(n)}$	0	0	0	0
$\bar{u}^{(m)}$	0	0	0	0

As required by Definition 1, the solutions defined by (69) and (77) constitute memory Nash equilibria of the game given by (54) and (55). By specifying particular threats and a particular bargaining model we have selected two from the infinitely many Nash equilibria of the game. Note, that any strategy of the form (69) with $\bar{u}^{(n)}$ replaced by an arbitrary u satisfying the enforceability conditions (46), (47) (there are continuum of such u) is also the Nash strategy. The same is true for the strategy (77) and a great number of other strategies having the similar form but differently formulated threats.

6. CONCLUSIONS

An equilibrium solution concept for dynamic games has been proposed where players can communicate with one another but cannot make binding contracts. In such games, unlike the static problems without contract possibilities, cooperation between players becomes possible due to the realization that agreements negotiated by the players can be enforced by suitably-defined strategies. The definition presented in Section 3 combines the dynamic programming, the theory of bargaining and the notion of enforceable agreements to produce a class of cooperative solutions defined in the form of memory Nash equilibria. A particular member of this class can be obtained by specifying players' disagreement policies, and a bargaining model providing the pattern for negotiation. In Section 4 we have discussed the most obvious choices of disagreement policies and bargaining models, and provided a sufficient condition for equilibrium in the case when the negotiation scheme proposed originally by Raiffa is selected as the bargaining model. The theoretical considerations have been illustrated by the example given in Section 5.

The proposed definition of dynamic equilibrium has several interesting generalizations. The formulation of threats used as the reference point in bargaining can be different from the formulation of threats used to prevent the violation of a negotiated agreement in the course of the game, that is d in the definition of the set $A_t(x_t, d^{t+1})$ can be replaced by d' different from the second argument of the bargaining model $F(G_t, d^t)$. The bargaining need not to take place at every stage of the game as required by Definition 1, instead the players can negotiate agreements for longer periods of time, and reopen negotiations every $k > 1$ stages, or if some more general conditions are satisfied. The retaliation for violating an agreement in the course of the game need not last until the end of the game, ruling out any possibility of reconciliation and return to the cooperative mood of play, but can be limited to fewer stages. Finally, the multiperson and stochastic games seem to be the promising field of further research.

REFERENCES

Aumann, R. J. (1959). Acceptable points in general cooperative *n*-person games. In A. W. Tucker and R. D. Luce (Eds), *Contributions to the Theory of Games*, Vol. IV. Princeton University Press, Princeton, pp. 287–332.

Başar, T. (1974). A counterexample in linear-quadratic games: existence of nonlinear Nash solutions. *J. Optimiz. Theory Appl.*, **14**, 425.

Başar, T. (1976). On the uniqueness of the Nash solution in linear-quadratic differential games. *Int. J. Game Theory*, **5**, 65.

Başar, T. and H. Selbuz (1979). Closed-loop Stackelberg strategies with applications in the optimal control of multilevel system. *IEEE Trans Aut. Control*, **AC-24**, 166.

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *Rev. Economic Studies*, **38**, 1.

Friedman, J. W. (1977). *Oligopoly and the Theory of Games.* North-Holland, Amsterdam.

Luce, R. D. and H. Raiffa (1957). *Games and Decisions: Introduction and Critical Survey.* John Wiley, New York.

Nash, J. F. (1950). The bargaining problem. *Econometrica*, **18**, 155.

Nash, J. F. (1953). Two-person, cooperative games. *Econometrica*, **21**, 128.

Owen, G. (1968). *Game Theory.* Saunders, Philadelphia.

Raiffa, H. 1953). Arbitration schemes for generalized two-person games. In. H. W. Kuhn and A. W. Tucker (Eds), *Contributions to the Theory of Games*, Vol. II. Princeton University Press, Princeton, pp. 361–387.

Rosenthal, R. W. (1976). An arbitration model for normal form games. *Math. Oper. Res.*, **1**, 82.

Roth, A. E. (1979). *Axiomatic Models of Bargaining.* Springer, New York.

Starr, A. W. and Y. C. Ho (1969). Nonzero-sum differential games. *J. Optimiz. Theory Appl.*, **3**, 184.

Tołwiński, B. (1979). Closed-loop Stackelberg solution and threats in dynamic games. *Proceedings 9th IFIP Conf. on Optimization Techniques*, Warsaw.

Tołwiński, B. (1980). Stackelberg solution of dynamic game with constraints. *Proceedings 19th IEEE CDC Conf.*, Albuquerque.

Tołwiński, B. (1981). Closed-loop Stackelberg solution to multistage linear-quadratic game. *J. Optimiz. Theory Appl.*, **34**, 485.