Games and Economic Behavior 71 (2011) 487–502

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Equilibrium play in matches: Binary Markov games ☆

Mark Walker, John Wooders *, Rabah Amir

Department of Economics, University of Arizona, Tucson, AZ 85721, USA

ARTICLE INFO

Article history:
Received 30 May 2009
Available online 12 May 2010

JEL classification:
C72
C73

Keywords:
Stochastic games
Minimax
Strictly competitive games
Game theory and sports
Tennis

ABSTRACT

We study two-person extensive form games, or "matches," in which the only possible outcomes (if the game terminates) are that one player or the other is declared the winner. The winner of the match is determined by the winning of points, in "point games." We call these matches *binary Markov games*. We show that if a simple monotonicity condition is satisfied, then (a) it is a Nash equilibrium of the match for the players, at each point, to play a Nash equilibrium of the point game; (b) it is a minimax behavior strategy in the match for a player to play minimax in each point game; and (c) when the point games all have unique Nash equilibria, the only Nash equilibrium of the binary Markov game consists of minimax play at each point. An application to tennis is provided.

1. Introduction

In many games, players are advised to play the same way regardless of the score. In tennis, for example, players are often advised by their coaches to play each point the same, whether the point is a match point or the opening point of the match. In poker it is a common belief that a player should play the same whether he is winning or losing.[1] This notion that one should play in a way that ignores the score stands in sharp contrast to the alternative view, also commonly espoused, that one should play differently on "big points," *i.e.*, in situations that are "more important." We establish in this paper that for a certain class of games, which we call *binary Markov games*, equilibrium play is indeed independent of the game's score.

We consider two-player games — we call them "matches" — which are composed of points, and in which (a) the history of points won can be summarized by a score, or state variable; (b) when the players compete with one another to win points, they do so via a "point game" which may depend upon the current score or state of the match; and (c) each player cares only about whether he wins or loses the match — *i.e.*, about whether the match terminates in a winning state for himself (and thus a loss for his opponent), or vice versa. We use the terminology "match" and "point game" in order to distinguish the overall game from the games in which the players compete for points.

Many real-life games, such as tennis, fit the description we have just given. It is useful, however, to begin by describing a much simpler example: two players play "matching pennies" repeatedly against one another; the winner of each matching pennies game wins a point; and the first player to be ahead by two points wins the overall game, or match. This game has only five potential scores, −2, −1, 0, +1, and +2, where the score +2 means Player A is ahead by two points (and has thus

☆ This research was partially supported by NSF grant #SBR-0099353. We are grateful to three referees and an advisory editor for detailed and insightful comments.

* Corresponding author.
E-mail addresses: mwalker@arizona.edu (M. Walker), jwooders@eller.arizona.edu (J. Wooders), ramir@eller.arizona.edu (R. Amir).

[1] But only in cash games, not in tournaments.

0899-8256/$ – see front matter
doi:10.1016/j.geb.2010.04.011

(a)

Even Points

		Column H	Column T	Row Minimax
Row	H	1	0	1/2
	T	0	1	1/2
Col Minimax:		1/2	1/2	

Value = 1/2 to both

(b)

Odd Points

		Column H	Column T	Row Minimax
Row	H	1	2/3	2/3
	T	1/3	1	1/3
Col Minimax:		1/3	2/3	

Value = 7/9 to Row
Value = 2/9 to Column

Fig. 1. Outcomes (cell entries) are the probability that row wins the point.

won the game, *i.e.*, the match), the score -2 means Player A has lost the match, the score 0 means the match is tied (this is the score when the match begins), and the scores $+1$ and -1 mean, respectively, that Player A is ahead by one point and that Player A is behind by one point. In this game, like all the games we consider, the players are interested in winning points only as a means to winning the match.

Now let's change the example slightly. Suppose that, as before, when the score is tied (*i.e.*, when the match is in state 0) the players play the conventional matching pennies game: let's say Row wins the point when the coins match, and Column wins the point when the coins don't match, as depicted in Fig. 1a. But when the score is not tied, and the match has not yet ended (*i.e.*, when the state is "odd," either $+1$ or -1), a slightly different matching-pennies game determines which player wins the current point, namely the game depicted in Fig. 1b. In this game the players still choose Heads or Tails, and Row still wins if their coins match. But if the coins don't match, then Nature randomly determines which player wins the point, and the player who chose Heads has a $2/3$ chance of winning, the player who chose Tails only a $1/3$ chance.

There are several things worth noting in this new match. First, just as in the conventional matching pennies game, the new "odd-state" point game has a unique equilibrium, which is in mixed strategies. But the equilibrium (and minimax) mixtures are different in the odd-state point game than in the conventional game: here the Row player plays Heads with mixture probability $2/3$ instead of $1/2$, and the Column player plays Heads with mixture probability $1/3$. The value of the odd-state point game to the Row player is $7/9$ (Row's probability of winning the point), and the game's value to the Column player is $2/9$. Thus, by always playing his minimax mixture in the current point game, the Row player can assure himself a probability $1/2$ of winning any point played when the score is even, and a $7/9$ probability of winning any point played when the score is not even. Similarly, the Column player, by always playing his point-game minimax strategy, can assure himself a $1/2$ probability of winning even points and a $2/9$ probability of winning odd points.

It is easy to verify that if the players always play their minimax mixtures in each point game, then the Row player will win the match with probability $7/9$ if the match is tied, with probability $77/81$ if he is ahead by a point, and with probability $49/81$ if he is behind by a point. Indeed, the Row player can assure himself of at least these probabilities of winning, no matter what his opponent does, by always playing his minimax mixture, and the parallel statement can be made for the Column player. Note that this "minimax play" prescription says that one's play should not depend on the match's score, except to the extent that the point game depends on the score: each player should play the same whether he is ahead in the match or behind (*viz.*, Row should mix $2/3$ on Heads, Column should mix $1/3$ on Heads).

It seems clear, at least intuitively, that such play is a Nash equilibrium in the match, and that a minimax strategy for playing the match is to play one's minimax mixture in every state of the match — and perhaps even that this is the *only* equilibrium in the match. Proving these propositions in general for binary Markov games, however, turns out to be non-trivial. This is largely because the match is not a game with a finite horizon: for example, if neither player in the example we have just described is ever ahead by two points, the match continues indefinitely.

We will identify a general class of games like the one above, in which we will show that equilibrium play at any moment is dictated only by the point game currently being played — play is otherwise independent of the history of points that have been won or actions that have been taken, and in particular it is independent of the score, except to the extent that the current point game depends on the score.

The results we obtain are for matches in which every point game is a strictly competitive win–loss game (*i.e.*, in each point game either one player or the other wins a single point).[2] Our first result, an Equilibrium Theorem, establishes that

[2] Wooders and Shachat (2001) also obtain results on equilibrium and minimax play in sequential play of stage games where at each stage one player wins and the other loses. Their stage games can thus be interpreted as contests for points. But in Wooders and Shachat (a) it is always strictly better for a player to win more points, and (b) the "match" has a known, finite length. In contrast, we assume here that (a) each player cares only about whether he wins or loses the match, and (b) the length of the match is allowed to be indefinite and infinite.

if the overall game satisfies a simple monotonicity condition, then there is an equilibrium of the match in which play is the same whenever the same point game arises (*i.e.*, play does not otherwise depend upon the score, or upon the history of points won or actions taken). Moreover, the equilibrium has a simple structure: at each score the players play the Nash equilibrium of the associated point game, in which winning the point is assigned a utility of one and losing the point is assigned a utility of zero.

Our second result, a Minimax Theorem, establishes that it is a minimax behavior strategy in the match for a player to play a minimax strategy in each point game: the match has a value for each player (*viz.*, his probability of winning the match under equilibrium play) and a player can *unilaterally* assure himself of at least this probability of winning the match, no matter how his opponent plays, by simply playing a minimax strategy in each point game. This result provides a rationale for the players to play minimax (and Nash equilibrium) in each point game, even if the usual behavioral assumptions underlying Nash equilibrium are not satisfied.

Our third result, a Uniqueness Theorem, establishes that if each point game has a unique minimax strategy for each player, as in the example above, then the only equilibrium in the match is for each player to always play his minimax strategy in each point game.

Our results apply to many real life games in which players compete for points and in which, at each score, either one player or the other wins the next point. We consider one application: in Section 8 we develop a model of a tennis match, in which there may be many different point games, or only a few. For example, which of the point games is currently in force may depend upon which player has the serve and upon which court — ad or deuce — he is serving from, but otherwise be independent of the score. We show that our game-theoretic model of tennis satisfies our monotonicity condition and hence that our results apply to tennis. Thus, despite the fact that tennis has an extremely complicated scoring rule,[3] equilibrium (and indeed minimax) play in a tennis match consists simply of Nash equilibrium play of the point game associated with the current score. This result provides the theoretical foundations for using field data from actual tennis matches to test the theory of mixed-strategy Nash equilibrium, as was done in Walker and Wooders (2001).

The remainder of the paper is organized as follows: In Section 2 we provide a formal definition of binary Markov games and three examples. In Section 3 we describe the relation between binary Markov games and previous research. In Sections 4, 5, and 6 we present the Equilibrium, Minimax, and Uniqueness Theorems for binary Markov games. In Section 7 we provide examples which show how our results depend upon both the monotonicity condition and the “binary” character of the point games (*i.e.*, that each has only two outcomes). Section 8 shows how our results apply to the game of tennis.

2. Binary Markov games

We begin by formalizing the kind of matches we will study — a special class of stochastic games that we refer to as *binary Markov games*. A binary Markov game has two elements, a *binary scoring rule*, and a collection of *point games*. We first describe the scoring rule, and then the point games. We will continue to use the term “match” as an informal synonym for a binary Markov game.

2.1. Binary scoring rule

A **binary scoring rule** consists of a finite set S of **states** and two **transition functions**, each of which maps S into S. The states represent, in a generalized sense, the possible scores in the match.[4] From every state, only two transitions to other states are possible: if the current state is s and Player A *wins* the current point, then we say the subsequent state is s_+; if instead A *loses* the point, the next state is s_-. The set S is assumed to include two **absorbing states**, ω_A (interpreted as the state in which A has won the match) and ω_B (the state in which B has won). Thus, we assume that $(\omega_A)_+ = (\omega_A)_- = \omega_A$ and $(\omega_B)_+ = (\omega_B)_- = \omega_B$.

Example 1. (*A best-two-out-of-three series of contests, such as a three-game playoff series*: *the winner of the match is the first player to win two contests.*) If we treat each contest as a point, then we can say that the winner of the match is the first player to win two points. In addition to the states ω_A and ω_B, there are four other states:

- 0: the initial state, in which no points have yet been played;
- A: Player A is ahead — he has won the first point;
- B: Player B is ahead — he has won the first point;
- T: the score is tied (two points have been played, and each player has won one of the points).

The transition functions s_+ and s_- are given by $0_+ = A$, $0_- = B$, $A_- = B_+ = T$, $A_+ = T_+ = \omega_A$, and $B_- = T_- = \omega_B$.

---

[3] The scoring rule is described in Section 8.

[4] In order to interpret a particular game or sport as a binary Markov game, it is often necessary that the set S distinguish scoring histories more finely than in the “score” as usually understood. For example, in sports such as squash and volleyball, it is necessary that the state include, in addition to the numerical score, a specification of which player has the serve.

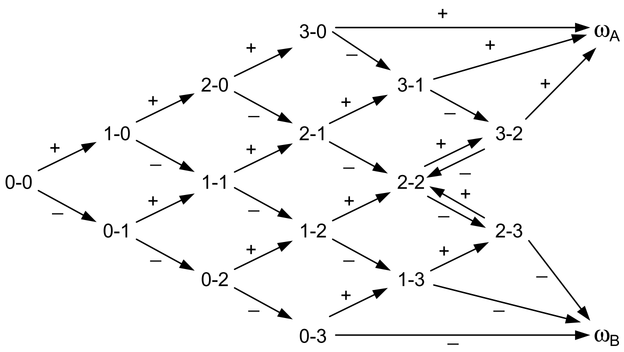

Fig. 2. The scoring rule for a "game" in tennis.

Example 2. (*A simplified version of tennis*: *the winner of the match is the first player to be ahead by two points.*) The score is said to be "deuce" when both players have won the same number of points; the deuce state is denoted 0, and it is the state in which the game begins. The player who wins a point played at deuce is then said to "have the advantage" until the subsequent point is completed. Let A and B denote the states in which Player A or Player B has the advantage. Clearly, if a player who has the advantage wins the next point then he wins the match, and if he loses the point then the score reverts to deuce. Thus, there are five states in all — ω_A, ω_B, 0, A, and B — and the transition functions are given by $0_+ = A$, $0_- = B$, $A_+ = \omega_A$, $B_- = \omega_B$, and $A_- = B_+ = 0$. This is exactly the same scoring rule as the one described above, in Section 1.

Note that there is an important difference between Examples 1 and 2. The game in Example 1 can last at most three points and then it will be finished: after three points have been played, it will be in either the state ω_A or the state ω_B. But the game in Example 2 can continue for an indefinite number of points, and there is the possibility that it never ends — *i.e.*, that it never enters either of the states ω_A or ω_B.

Example 3. (*A single "game" in a tennis match*: *the winner is the first player to win at least four points and to simultaneously have won at least two points more than his opponent.*) The states are the five states in Example 2, as well as the additional "transitory" states that correspond to the possible scores when neither player has yet won four points. (These states are transitory in the sense that once the game enters such a state, it must immediately exit from the state, never to return to that state.) It is easy to verify that there are twelve such transitory states; see Fig. 2. If the game has not terminated by the time six points have been played (in which case each player has won exactly three points), then the situation is exactly equivalent to a score of 2–2, referred to as "deuce," and from that point forward the game is identical to the game in Example 2.

The binary scoring rule describes how the various possible states, or situations in the match, are linked to one another — *i.e.*, it tells us what will happen if Player A or Player B wins a given point — but it tells us nothing about *how* the points are won or lost, *i.e.*, about how the players' actions determine the winning or losing of points. We assume that the winner of each point is determined by the outcome of a normal form game between the two players, and we assume that the details of this *point game* depend only upon the current state of the match.

Thus, for each state $s \in S$, we define the **point game associated with *s*** to be a normal form game G_s, with finite **action sets** $A(s)$ and $B(s)$ for Players A and B, and with **point game payoff functions** $\pi_{si} : A(s) \times B(s) \to [0, 1]$ for each player $i \in \{A, B\}$. The payoff $\pi_{si}(a, b)$ is the probability, when actions a and b are chosen while in state s, that player i will win the point. (Thus, the functions π_{si} are payoff functions in name only and are not directly related to the payoffs in the match: they merely determine, along with the functions s_+ and s_-, the transitions between states.) We require that every point be won by one of the players, *i.e.*, that $\pi_{sA}(\cdot) + \pi_{sB}(\cdot) \equiv 1$. Each game G_s is therefore a constant-sum game and hence has a *value*, which we denote by $v_A(s)$ for Player A and $v_B(s)$ for Player B. For completeness we assume that $\pi_{\omega_A A}(\cdot,\cdot) \equiv \pi_{\omega_B B}(\cdot,\cdot) \equiv 1$ and $\pi_{\omega_A B}(\cdot,\cdot) \equiv \pi_{\omega_B A}(\cdot,\cdot) \equiv 0$.

If each player plays a minimax strategy when in state s, then $v_A(s)$ and $v_B(s)$ are the transition probabilities, *i.e.*, the probabilities of moving to state s_+ or s_-. Of course, the players need not play their minimax strategies. For any actions a and b that they choose, the payoff numbers $\pi_{sA}(a, b)$ and $\pi_{sB}(a, b)$ are the respective transition probabilities of moving from state s to state s_+ or state s_-. In general, when actions a and b are chosen, we denote the probability of moving from state s to *any* other state s' by $P_{ss'}(a, b)$. In other words,

$$P_{ss'}(a, b) = \begin{cases} \pi_{sA}(a, b), & \text{if } s' = s_+, \\ \pi_{sB}(a, b), & \text{if } s' = s_-, \\ 0, & \text{otherwise.} \end{cases}$$

Note that the functions π_{si} are not the payoff functions for the binary Markov game — *i.e.*, for the match. Rather, the scoring rule and the point games (including their payoff functions) merely specify how the players' actions determine the transitions from state to state. Thus, the functions $P_{ss'}(a,b)$ are the transition functions of a stochastic game.[5] It may be of interest to point out that one could equivalently treat the scoring rule and the point games together as a *game form* that describes how sequences of actions (stochastically) determine outcomes, *i.e.*, sequences of states.

2.2. Strategies and payoffs in binary Markov games

It remains to define the players' payoff functions for the match. A match has three possible outcomes: it can end in either of the two terminal states ω_A or ω_B, or it can continue forever, never entering either of the terminal states. We assume that a player receives a positive payoff (normalized to 1) for winning, and a zero payoff otherwise.[6]

In order to define the players' payoff functions for the match, we need to establish what we mean by a *strategy* for playing the match: we define a *behavior strategy* for a player, which specifies what action the player will take after any possible *history of play* in the match. We say that a **history at time** t consists of the current state, s_t, as well as all states and actions that have occurred prior to t. In other words, a history at t is a $(3t+1)$-tuple $h_t = (s_0,\dots,s_t;a_0,\dots,a_{t-1};b_0,\dots,b_{t-1})$, where, for each $k \leqslant t$, s_k is the state at time k, and where (for $k<t$) $a_k \in A(s_k)$ and $b_k \in B(s_k)$ are the actions the players chose at time k. Denote by H_t the set of all possible histories[7] at time t, and let H denote the set of all possible histories: $H = \bigcup_{t=0}^{\infty} H_t$. A **behavior strategy** for Player A is a function α which, for every history $h_t \in H$, prescribes a probability distribution (*i.e.*, a mixture) over the action set $A(s_t)$. The mixture probability that α assigns to an action $a \in A(s_t)$ is denoted by $\alpha(a|h_t)$. A behavior strategy α is **stationary** if its prescription depends only upon the current state — *i.e.*, if whenever two histories h_t and $h_{t'}$ satisfy $s_t = s_{t'}$, then $\alpha(h_t) = \alpha(h_{t'})$. When α is a stationary strategy, we often write $\alpha(s_t)$ instead of $\alpha(h_t)$, and $\alpha(a|s_t)$ instead of $\alpha(a|h_t)$. Behavior strategies β for Player B, and the notations $\beta(h_t)$, $\beta(b|h_t)$, $\beta(s_t)$ and $\beta(b|s_t)$, are defined analogously. If h_t is a history at time t, and if a and b are actions in $A(s_t)$ and $B(s_t)$, then $h_t + (s_{t+1}, a, b)$ will denote the history at $t+1$ in which the actions a and b were taken at t and then the state s_{t+1} occurred.

For every pair (α,β) of behavior strategies, we denote by $p_s^{\alpha,\beta}(h_t)$ the probability that the history at time t will be h_t if the initial state is s and Players A and B follow the behavior strategies α and β. This probability is defined recursively:

$$p_s^{\alpha,\beta}(h_0) := \begin{cases} 1 & \text{if } s = s_0, \\ 0 & \text{if } s \neq s_0, \end{cases}$$

and for $h_{t+1} = h_t + (s_{t+1}, a_t, b_t)$,

$$p_s^{\alpha,\beta}(h_{t+1}) := p_s^{\alpha,\beta}(h_t)\alpha(a_t|h_t)\beta(b_t|h_t)P_{s_t s_{t+1}}(a_t,b_t).$$

The payoff a player receives when the players employ strategies α and β is simply the resulting probability the player will win the match. We will generally denote these "winning probabilities" by W_A and W_B. Evaluating how these probabilities depend on the players' strategies α and β will require dynamic programming arguments in which we will also have to consider how the probabilities depend upon the current state s. Thus, let $W_A(s,\alpha,\beta)$ denote the probability (at $t=0$) that Player A will eventually win the match if the *initial* state is s and if the players follow the behavior strategies α and β; *i.e.*,

$$\forall s \in S: \quad W_A(s,\alpha,\beta) = \lim_{t\to\infty} \sum_{\{h_t \in H_t | s_t = \omega_A\}} p_s^{\alpha,\beta}(h_t).$$

(Note that the limit in the above expression exists, because the sequence is increasing in t and is bounded above by 1.) The function $W_B(s,\alpha,\beta)$ is defined similarly. We will use the function $W_A(\cdot)$ to analyze Player A's best response function; we will rarely need to make explicit use of the function $W_B(\cdot)$, because the analysis for Player A can be applied directly to Player B.

For non-terminating histories — those in which play never enters the absorbing class $\{\omega_A,\omega_B\}$ — there is no winner of the match, and each player's payoff according to the above limit is zero. Consequently the game is not a constant-sum game, even though for all histories that enter the absorbing class $\{\omega_A,\omega_B\}$ the sum of the players' payoffs is 1.[8,9]

---

[5] The stochastic games literature often adds the assumption that from every state s there is a positive probability that the next transition will end the game — *i.e.*, will be to a terminal state. This is clearly not the case for binary Markov games, where the binary nature of the transitions rules this out.

[6] If we instead define Player A's payoff as 1 in state ω_A, -1 in state ω_B, and zero if the game never terminates, and the opposite for Player B, then the match is a zero-sum recursive game, exactly the case studied by Everett (1957). However, in that case the payoffs are not binary, and therefore it is not included in the definition of a binary Markov game. Also see footnote 9 below.

[7] We include only histories that are consistent with the scoring rule, *i.e.*, for which s_{k+1} is always either $(s_k)_+$ or $(s_k)_-$. Note, however, that we allow s_0 to be any $s \in S$; *i.e.*, $H_0 = S$, so that formally the match can begin in any state. This is important for the dynamic programming arguments we will employ.

[8] If a binary Markov game satisfies a monotonicity condition to be introduced below, and if both players play minimax strategies in every point game, then play will enter the absorbing class with probability one, as shown in Lemma 2, below. A referee has suggested that these games are therefore "almost zero-sum."

[9] It is an open question whether our results go through if a binary Markov game is altered to make it a constant-sum game, as described for example in footnote 6. Because the payoffs are no longer binary, the proof we provide below for Lemma 1 no longer goes through. We have been unable to establish a proof or a counterexample for this case.

3. Stochastic and recursive games

Binary Markov games are related to several kinds of dynamic, multi-stage games that have been studied extensively. Shapley (1953) introduced stochastic games, describing them as games in which "play proceeds by steps from position to position, according to transition probabilities controlled jointly by the two players." Everett (1957) introduced recursive games — stochastic games that have absorbing states and in which payoffs occur only when an absorbing state is reached. A binary Markov game is thus a recursive game in which both the state transitions and the payoffs are binary.[10]

For undiscounted stochastic games with finite state and action spaces, Mertens and Neyman (1981) established the existence of a value and of ε-minimax strategies in the zero-sum case, and Vieille (2000a, 2000b) established the existence of an ε-equilibrium in the nonzero-sum case. In each case, the strategies may be history-dependent. Blackwell and Ferguson's (1968) insightful Big Match, a zero-sum game, shows that in general neither the history dependence nor the ε-character of the strategies can be improved upon: the Big Match has neither an exact equilibrium nor a stationary ε-equilibrium.

A binary Markov game is a special case of the two settings in Vieille (2000a, 2000b). Therefore the existence of an ε-equilibrium in history-dependent strategies follows directly from his work. But we wish to obtain substantially sharper results than this, results that emerge from the added structure of the binary transitions and payoffs of binary Markov games. We will establish that a *specific* strategy, one that is intuitively appealing and theoretically important — namely, the stationary strategy of playing minimax at each stage of the game, independently of history — is a minimax strategy and an equilibrium strategy (with no ε-qualification) in any binary Markov game, even though these are not, strictly speaking, zero-sum games. Moreover, if the point games have unique equilibria, then the unique equilibrium of the binary Markov game is for each player to play this strategy. Note that the minimax result provides a rationale for the simple strategy of always playing minimax in every point game, without the stronger assumptions often used to justify Nash equilibrium.

Some research has focused on the existence of ε-equilibrium in Markov-stationary strategies for undiscounted stochastic games.[11] Existence has been established in several classes of games, but the results suggest that the scope for stationary ε-equilibria in general undiscounted stochastic games is quite limited.

Milnor and Shapley (1957) introduced games of survival, a special class of recursive games in which the state space is linearly ordered, transitions are governed by a single matrix game in which each of the outcomes is a shift operator on the state space, and the largest and smallest states are the absorbing states; in one of the absorbing states Player A is the winner and in the other Player B is the winner. Unlike binary Markov games, the only stochastic feature of a game of survival is not in the game itself, but arises from the mixed strategies employed by the players. While binary Markov games are similar to games of survival, neither is a special case of the other: transitions need not be binary in a game of survival, and transition games (the point games) in a binary Markov game need not be the same in each state, nor are the states necessarily linearly ordered. Clearly, however, a game of survival in which the transitions *are* binary is a binary Markov game, and our results therefore apply to such games. These distinctions can be seen in the three examples in Section 2. In Example 2 the transition function is a binary shift operator on the linearly ordered state space. Thus, if the transitions in the three non-absorbing states are all determined by the same point game, then the example is a typical (albeit simple) game of survival. If the point games differ, the example remains a binary Markov game, and could be called a generalized game of survival. By contrast, in Examples 1 and 3 the state space is not linearly ordered, so these examples cannot be games of survival for any point games.

4. The Equilibrium Theorem

Our aim is to establish that equilibrium strategies for playing the match have certain characteristics. We begin by identifying an *Optimality Equation* which must be satisfied for any strategy that maximizes a player's state-contingent probabilities of winning the match. A dynamic programming argument establishes Lemma 1, which states that if a stationary strategy satisfies the Optimality Equation, then that strategy is a best response. It will then be straightforward to use Lemma 1 to establish the Equilibrium Theorem, which states that a strategy which prescribes minimax play in each state is a best response to such a strategy by the opposing player.

In Section 2 we defined $W_i(s,\alpha,\beta)$ as the probability that Player i will eventually win the match if the *initial state* is s and if the players follow the (not necessarily stationary) behavior strategies α and β. But when α and β *are* stationary, the winning probability must be the same from a given state whenever it is reached. Therefore, in the case of stationary α and β, the probabilities $W_i(s,\alpha,\beta)$ are not merely the probabilities Player i will win if the *initial* state is s, but they are Player i's *continuation payoffs*, the probability he will ultimately win the continuation game defined by having reached the state s.

When Player B is playing a stationary strategy β, we will use the notation $\bar{P}_{ss'}(a,\beta)$ for the transition probabilities available to Player A — *i.e.*, when the current state is s and Player A chooses action $a \in A(s)$, then $\bar{P}_{ss'}(a,\beta)$ is the probability that the next state will be s':

$$\forall s,\ s' \in S,\ a \in A(s):\quad \bar{P}_{ss'}(a,\beta) = \sum_{b \in B(s)} \beta(b|s) P_{ss'}(a,b).$$

[10] Surveys of stochastic games can be found in Mertens (2002), Vieille (2002), and Neyman and Sorin (2003).

[11] See, for example, Tijs and Vrieze (1986), Thuijsman and Vrieze (1991), and Flesch et al. (1996).

In Lemma 1 we establish a sufficient condition for a stationary strategy α to be a best response to a stationary strategy β. Clearly, a *necessary* condition that the probabilities $W_A(s,\alpha,\beta)$ must satisfy if α and β are both stationary and if α is a best response to β is the following **Optimality Equation**:

$$\forall s \in S: \quad W_A(s,\alpha,\beta) = \max_{a\in A(s)} \sum_{s'\in S} \bar{P}_{ss'}(a,\beta) W_A\bigl(s',\alpha,\beta\bigr). \tag{1}$$

Lemma 1 tells us that the necessary condition is sufficient as well: if β is a stationary strategy for Player B, and if a stationary strategy α for Player A generates a configuration of winning probabilities $W_A(s,\alpha,\beta)$ that satisfies the Optimality Equation, then α is a best response (among *all* of Player A's possible behavior strategies) to β.

Lemma 1. *Let α and β be stationary behavior strategies. If the probabilities $W_A(s,\alpha,\beta)$ satisfy the Optimality Equation, then α is a best response to β. That is, if*

$$\forall s \in S: \quad W_A(s,\alpha,\beta) = \max_{a\in A(s)} \sum_{s'\in S} \bar{P}_{ss'}(a,\beta) W_A\bigl(s',\alpha,\beta\bigr), \tag{2}$$

then $W_A(s,\alpha,\beta) \geqslant W_A(s,\alpha',\beta)$ for each state $s \in S$ and for every one of Player A's (not necessarily stationary) behavior strategies α'.

Proof. Assume that the probabilities $W_A(s,\alpha,\beta)$ satisfy the Optimality Equation, and let α' be an arbitrary strategy (not necessarily stationary) for Player A. By (2), for any time t and any history $h \in H_t$ we have

$$\begin{aligned} W_A(s_t,\alpha,\beta) &= \max_{a\in A(s_t)} \sum_{s'\in S} \sum_{b\in B(s_t)} \beta(b|s_t) P_{s_t s'}(a,b) W_A\bigl(s',\alpha,\beta\bigr) \\ &\geqslant \sum_{a\in A(s_t)} \alpha'(a|h) \sum_{s'\in S} \sum_{b\in B(s_t)} \beta(b|s_t) P_{s_t s'}(a,b) W_A\bigl(s',\alpha,\beta\bigr) \\ &= \sum_{s'\in S} \sum_{a\in A(s_t)} \sum_{b\in B(s_t)} \alpha'(a|h)\beta(b|s_t) P_{s_t s'}(a,b) W_A\bigl(s',\alpha,\beta\bigr). \end{aligned} \tag{3}$$

From any current state s, the next state will certainly satisfy either $s' = \omega_A$ or $s' \neq \omega_A$, and we can therefore rewrite (3) as

$$\begin{aligned} W_A(s_t,\alpha,\beta) \geqslant{} & \sum_{a\in A(s_t)} \sum_{b\in B(s_t)} \alpha'(a|h)\beta(b|s_t) P_{s_t \omega_A}(a,b) \\ & + \sum_{s'\neq\omega_A} \sum_{a\in A(s_t)} \sum_{b\in B(s_t)} \alpha'(a|h)\beta(b|s_t) P_{s_t s'}(a,b) W_A\bigl(s',\alpha,\beta\bigr). \end{aligned} \tag{4}$$

Let $W(s,\alpha';t,\alpha;\beta)$ denote the probability (at time zero, the beginning of the match) that Player A will eventually win the match, if the initial state is s, and if he follows α' through time $t-1$ and he then follows α (which is stationary) subsequently. For every $t \geqslant 1$ we of course have

$$W\bigl(s,\alpha';t,\alpha;\beta\bigr) = \sum_{\{h\in H_t | s_t=\omega_A\}} p_s^{\alpha',\beta}(h) + \sum_{\{h\in H_t | s_t\neq\omega_A\}} p_s^{\alpha',\beta}(h) W_A(s_t,\alpha,\beta). \tag{5}$$

For $t = 1$, Eq. (5) is

$$W\bigl(s,\alpha';1,\alpha;\beta\bigr) = \sum_{a\in A(s)} \sum_{b\in B(s)} \alpha'(a|s)\beta(b|s) P_{s\omega_A}(a,b) + \sum_{s'\neq\omega_A} \sum_{a\in A(s)} \sum_{b\in B(s)} \alpha'(a|s)\beta(b|s) P_{ss'}(a,b) W_A\bigl(s',\alpha,\beta\bigr),$$

and therefore, according to (4), we have $W_A(s,\alpha,\beta) \geqslant W(s,\alpha';1,\alpha;\beta)$, for each $s \in S$ – *i.e.*, the probability that Player A will eventually win if the initial state is s and he follows α is at least as great as the probability he would win if he instead followed α' for one period and then followed α subsequently.

We now show that $W(s,\alpha';t,\alpha;\beta) \geqslant W(s,\alpha';t+1,\alpha;\beta)$ for every $t \geqslant 1$. Replacing $W_A(s_t,\alpha,\beta)$ in (5) with the right-hand side of (4) we have

$$\begin{aligned} W\bigl(s,\alpha';t,\alpha;\beta\bigr) \geqslant{} & \sum_{\{h\in H_t | s_t=\omega_A\}} p_s^{\alpha',\beta}(h) + \sum_{\{h\in H_t | s_t\neq\omega_A\}} p_s^{\alpha',\beta}(h) \\ & \times \Biggl(\sum_{a\in A(s_t)} \sum_{b\in B(s_t)} \alpha'(a|h)\beta(b|s_t) P_{s_t\omega_A}(a,b) \\ & + \sum_{s_{t+1}\neq\omega_A} \sum_{a\in A(s_t)} \sum_{b\in B(s_t)} \alpha'(a|s_t)\beta(b|s_t) P_{s_t s_{t+1}}(a,b) W_A(s_{t+1},\alpha,\beta) \Biggr). \end{aligned} \tag{6}$$

The right-hand side of (6) can be rewritten as

$$\sum_{\{h\in H_t|s_t=\omega_A\}} p_s^{\alpha',\beta}(h) + \sum_{\{h\in H_t|s_t\neq\omega_A\}} p_s^{\alpha',\beta}(h) \sum_{a\in A(s_t)} \sum_{b\in B(s_t)} \alpha'(a|h)\beta(b|s_t)P_{s_t\omega_A}(a,b)$$
$$+ \sum_{\{h\in H_t|s_t\neq\omega_A\}} p_s^{\alpha',\beta}(h) \sum_{s_{t+1}\neq\omega_A} \sum_{a\in A(s_t)} \sum_{b\in B(s_t)} \alpha'(a|s_t)\beta(b|s_t)P_{s_ts_{t+1}}(a,b)W_A(s_{t+1},\alpha,\beta).$$

The first term in this sum is the probability, at time 0, that Player A will win by time t (*i.e.*, that $s_t=\omega_A$); the second term is the probability at time 0 that Player A will win *at* time $t+1$ (*i.e.*, that $s_t\neq\omega_A$ and $s_{t+1}=\omega_A$); and the last term is the probability at time 0 that Player A will win, but at a time later than $t+1$. Hence we have

$$W\big(s,\alpha';t,\alpha;\beta\big) \geqslant \sum_{\{h\in H_{t+1}|s_{t+1}=\omega_A\}} p_s^{\alpha',\beta}(h) + \sum_{\{h\in H_{t+1}|s_{t+1}\neq\omega_A\}} p_s^{\alpha',\beta}(h)W_A(s_{t+1},\alpha,\beta)$$
$$= W\big(s,\alpha';t+1,\alpha;\beta\big).$$

We have shown for an arbitrary behavior strategy α' that $W_A(s,\alpha,\beta)\geqslant W(s,\alpha';1,\alpha;\beta)$ and that $W(s,\alpha';t,\alpha;\beta)\geqslant W(s,\alpha';t+1,\alpha;\beta)$ for every t. Consequently, $W_A(s,\alpha,\beta)\geqslant W(s,\alpha';t,\alpha;\beta)$, for every $t\geqslant 1$. When combined with (5) this implies that for each t we have

$$W_A(s,\alpha,\beta) \geqslant \sum_{\{h\in H_t|s_t=\omega_A\}} p_s^{\alpha',\beta}(h). \tag{7}$$

Since $W_A(s,\alpha',\beta)$ is the limit of the right-hand side of (7) as t grows large, we have $W_A(s,\alpha,\beta)\geqslant W_A(s,\alpha',\beta)$. And since α' was an arbitrary behavior strategy and s an arbitrary initial state, α is therefore a best response to β. □

Lemma 1 provides a sufficient condition for ensuring that a stationary strategy α is a best response to a given stationary strategy β, but it does not *exhibit* a best response to any particular β, nor does it ensure that a best response even exists for any particular β. In the following Equilibrium Theorem we show that if β is a stationary strategy in which Player B always (in every state s) plays a minimax strategy in the point game G_s, then it is a best response for Player A to do the same. Thus, it is an equilibrium for each player to follow such a strategy.

We first reproduce the standard definition of a minimax strategy in a finite game, where for any finite set X we use ΔX to denote the set of probability distributions over X:

Definition. For each state $s\in S$, a **maximin strategy** for Player A in the point game G_s is a mixture $\alpha_s\in\Delta A(s)$ that satisfies

$$\alpha_s \in \arg\max_{\alpha'_s\in\Delta A(s)} \min_{\beta_s\in\Delta B(s)} \sum_{a\in A(s)} \sum_{b\in B(s)} \alpha'_s(a)\beta_s(b)\pi_{sA}(a,b).$$

Because each game G_s is a finite constant-sum game, a maximin strategy is also a **minimax strategy**, *i.e.*,

$$\alpha_s \in \arg\min_{\alpha'_s\in\Delta A(s)} \max_{\beta_s\in\Delta B(s)} \sum_{a\in A(s)} \sum_{b\in B(s)} \alpha'_s(a)\beta_s(b)\pi_{sB}(a,b).$$

We refer to a behavior strategy which prescribes minimax play in each state s as a **minimax-stationary strategy**.[12] Recall that for each state s, the value of the point game G_s to Player A is $v_A(s)$ and to Player B it is $v_B(s)$. Thus, if each player plays a minimax-stationary strategy, then the transition probabilities from a state s to the states s_+ and s_- are simply $v_A(s)$ and $v_B(s)$. For each state s, let $W_i^v(s)$ denote the probability that Player i will eventually win if these are the transition probabilities and if s is the current state. Then

$$\forall s\in S\colon\ W_A^v(s) = v_A(s)W_A^v(s_+) + v_B(s)W_A^v(s_-). \tag{8}$$

Note that $W_A^v(\omega_A)=1$ and $W_A^v(\omega_B)=0$.

In order to establish the Equilibrium Theorem, which states that it is an equilibrium in a binary Markov game for each player to play a minimax-stationary strategy, we must restrict slightly the class of binary Markov games we will consider. Consider a game, for example, in which, in some state s, a player's probability of winning the match is *reduced* if he wins the point game G_s. In such a game, the player will typically *not*, as part of a best response, play to win the point game G_s, and therefore he will not play a minimax strategy in G_s. The following straightforward condition rules out such pathologies.

[12] If a point game has more than one minimax strategy for a player, this definition allows for playing different minimax strategies in the same state at different times. Such a strategy is not, strictly speaking, stationary. Clearly this will not affect any of the arguments, and we will not explicitly make a distinction between minimax-stationary strategies and truly stationary strategies in which a player always plays minimax.

Definition. A binary Markov game satisfies the **Monotonicity Condition** if for each non-terminal state $s \in S$, $W_A^v(s_+) > W_A^v(s_-)$ and $W_B^v(s_-) > W_B^v(s_+)$.

This is an appealing condition that is likely to be satisfied by most binary Markov games one will encounter. It is satisfied in all the examples in this paper and for generalizations of the examples. As demonstrated in the application in Section 8, the condition can be verified in a given game via a system of linear equations that requires knowing only the values of the point games and how the states are linked to one another by the transition law.

Equilibrium Theorem. *If a binary Markov game satisfies the Monotonicity Condition, then against a minimax-stationary strategy β for Player B, any minimax-stationary strategy α is a best response for Player A; and against a minimax-stationary strategy α for Player A, any minimax-stationary strategy β is a best response for Player B. Thus, any pair of minimax-stationary strategies is a Nash equilibrium of the binary Markov game.*

Proof. According to Lemma 1, it will be sufficient to establish that the winning probabilities $W_A(s,\alpha,\beta)$ satisfy the Optimality Equation. Let $W^*(\cdot)$ be the function defined by the right-hand side of the Optimality Equation and by the function $W_A(s,\alpha,\beta)$, *i.e.*,

$$\forall s \in S: \quad W^*(s) := \max_{a\in A(s)} \sum_{s'\in S} \sum_{b\in B(s)} \beta(b|s) P_{ss'}(a,b) W_A\bigl(s',\alpha,\beta\bigr).$$

We must show that the two functions $W^*(\cdot)$ and $W_A(\cdot,\alpha,\beta)$ are identical. For each $s \in S$ we have

$$\begin{aligned} W^*(s) &= \max_{a\in A(s)} \sum_{b\in B(s)} \beta(b|s)\bigl[\pi_{sA}(a,b)W_A(s_+,\alpha,\beta) + \pi_{sB}(a,b)W_A(s_-,\alpha,\beta)\bigr] \\ &= \max_{a\in A(s)} \left[W_A(s_+,\alpha,\beta) \sum_{b\in B(s)} \beta(b|s)\pi_{sA}(a,b) + W_A(s_-,\alpha,\beta) \sum_{b\in B(s)} \beta(b|s)\pi_{sB}(a,b) \right]. \end{aligned}$$

The Monotonicity Condition ensures that $W_A(s_+,\alpha,\beta) > W_A(s_-,\alpha,\beta)$ for each state s, and therefore, since $\pi_{sA}(a,b) + \pi_{sB}(a,b) \equiv 1$, an action a maximizes the expression in brackets above if and only if a maximizes

$$\sum_{b\in B(s)} \beta(b|s)\pi_{sA}(a,b). \tag{9}$$

Since for each s the mixtures $\alpha(\cdot|s)$ and $\beta(\cdot|s)$ are minimax strategies in the point game G_s, every action in the support of the mixture $\alpha(\cdot|s)$ must maximize (9). Thus, for each $a \in \operatorname{supp} \alpha(\cdot|s)$,

$$W^*(s) = W_A(s_+,\alpha,\beta) \sum_{b\in B(s)} \beta(b|s)\pi_{sA}(a,b) + W_A(s_-,\alpha,\beta) \sum_{b\in B(s)} \beta(b|s)\pi_{sB}(a,b),$$

and therefore also

$$\begin{aligned} W^*(s) &= W_A(s_+,\alpha,\beta) \sum_{a\in A(s)} \alpha(a|s) \sum_{b\in B(s)} \beta(b|s)\pi_{sA}(a,b) + W_A(s_-,\alpha,\beta) \sum_{a\in A(s)} \alpha(a|s) \sum_{b\in B(s)} \beta(b|s)\pi_{sB}(a,b) \\ &= W_A(s_+,\alpha,\beta)v_A(s) + W_A(s_-,\alpha,\beta)v_B(s) \\ &= W_A^v(s), \quad \text{according to (8)} \\ &= W_A(s,\alpha,\beta). \qquad \square \end{aligned}$$

5. The Minimax Theorem

The Equilibrium Theorem tells us that "always playing minimax" is a best way to play in the match if one's opponent is playing that way. We show here that it is actually a *minimax* behavior strategy in the match to always play minimax: by playing minimax for every point, a player assures himself that his probability of winning the match will be at least as great as in equilibrium, no matter how his opponent plays.

We first provide definitions of maximin and minimax behavior strategies for playing the match.

Definition. A behavior strategy $\hat{\alpha}$ is a **maximin behavior strategy** for Player A in the binary Markov game if

$$\hat{\alpha} \in \arg\max_{\alpha} \min_{\beta} W_A(s,\alpha,\beta), \quad \text{for every } s \in S, \tag{10}$$

and $\hat{\alpha}$ is a **minimax behavior strategy** for Player A if

$$\hat{\alpha} \in \arg\min_{\alpha} \max_{\beta} W_B(s,\alpha,\beta), \quad \text{for every } s \in S. \tag{11}$$

We prove a Minimax Theorem for binary Markov games, which establishes that $\max_\alpha \min_\beta W_A(s,\alpha,\beta)$ and $\min_\beta \max_\alpha W_A(s,\alpha,\beta)$ both exist and that they are both equal to $W_A^v(s)$, *i.e.*, to Player A's probability of winning the match from state s when both players play (any) minimax-stationary strategies. Thus, $W_A^v(s)$ is the *value* of the binary Markov game for Player A, if the game is begun in state s. Furthermore, every minimax-stationary strategy is shown to be a minimax behavior strategy. The parallel results hold for Player B, of course, and the value of the game to him, starting from state s, is $W_B^v(s)$. Thus, by simply adopting a minimax-stationary behavior strategy, a player can guarantee himself a probability of at least $W_A^v(s)$ or $W_B^v(s)$ of winning the match, starting from a given state s.

Minimax Theorem for Binary Markov Games. *Suppose that a binary Markov game satisfies the Monotonicity Condition and that, for each non-terminal state s, we have* $0 < v_A(s) < 1$. *Then for every* $s \in S$, $W_A^v(s)$ *is the value of the game to Player A when the game is begun in state s,* i.e.,

$$\max_\alpha \min_\beta W_A(s,\alpha,\beta) = \min_\beta \max_\alpha W_A(s,\alpha,\beta) = W_A^v(s),$$

and similarly for Player B. Moreover, every minimax-stationary strategy for either player is both a minimax behavior strategy and a maximin behavior strategy.

The theorem's proof is a straightforward application of the following lemma.

Lemma 2. *Suppose that a binary Markov game satisfies the Monotonicity Condition, and suppose that at each non-terminal state s the value of the point game to Player A is strictly positive:* $v_A(s) > 0$. *If* $\hat{\alpha}$ *is a minimax-stationary strategy for Player A, then for any initial state s and for any behavior strategy* β *for Player B we have* $W_A(s,\hat{\alpha},\beta) + W_B(s,\hat{\alpha},\beta) = 1$, i.e., $\Pr(s_t \in \{\omega_A,\omega_B\}$ *for some* $t) = 1$.

Proof. Let K denote the number of non-terminal states, *i.e.*, $K := |S \setminus \{\omega_A,\omega_B\}|$; and for any state s and any non-negative integer k, let s^k denote the state reached when Player A wins k consecutive points beginning at state s – *i.e.*,

$$s^k := s\underbrace{+\cdots+}_{k\,\text{times}}.$$

We establish first that if Player A wins K consecutive points (beginning from any non-terminal state s) then he wins the match (the binary Markov game) – *i.e.*, $s^K = \omega_A$. Suppose to the contrary that $s^K \neq \omega_A$ for some non-terminal state s. Then, since the state ω_A is absorbing, it cannot be the case that $s^k = \omega_A$ for some $k < K$. Nor can it be the case that $s^k = \omega_B$ for some $k \leqslant K$, for then there would have to be a *first* state among $s, s^1, \ldots, s^K$ which is ω_B, say s^k, and the Monotonicity Condition would then yield $W_A^v(s^{k-1}) < W_A^v(\omega_B) = 0$, which is impossible. Hence, since none of the states s^k is terminal for $k \leqslant K$, the Monotonicity Condition yields

$$W_A^v(s) < W_A^v(s^1) < \cdots < W_A^v(s^K),$$

and therefore, in the course of winning K consecutive points without winning the match, Player A must visit $K+1$ *distinct* non-terminal states. This is a contradiction, since there are only K non-terminal states; hence, A wins the match if he wins K consecutive points.

In order to complete the proof, assume that the initial state, s_0, is a non-terminal state (otherwise the Lemma is trivially true). Let $q = \min\{v_A(s) \mid s \in S \setminus \{\omega_A,\omega_B\}\} > 0$. Then the probability Player A will have won by period K is $\Pr(s_K = \omega_A) \geqslant q^K$ since, when following a minimax-stationary strategy, Player A's probability of winning is at least q on each point, and since Player A wins the match if he wins K consecutive points. Clearly, $\Pr(s_K \in \{\omega_A,\omega_B\}) \geqslant q^K$ as well. Hence $\Pr(s_K \notin \{\omega_A,\omega_B\}) \leqslant 1 - q^K$. Furthermore, for each positive integer n, $\Pr(s_{nK} \notin \{\omega_A,\omega_B\}) \leqslant (1-q^K)^n$. It follows that $\Pr(\forall n\colon\ s_{nK} \notin \{\omega_A,\omega_B\}) = \lim_{n\to\infty}(1-q^K)^n = 0$. $\square$

Proof of the BMG Minimax Theorem. Let $\hat{\alpha}$ and $\hat{\beta}$ be minimax-stationary strategies for Players A and B. The Equilibrium Theorem guarantees that $\hat{\alpha}$ is a best response to $\hat{\beta}$, so we have $W_A(s,\alpha,\hat{\beta}) \leqslant W_A(s,\hat{\alpha},\hat{\beta})$ for every behavior strategy α. Similarly, $\hat{\beta}$ is a best response to $\hat{\alpha}$, so we have $W_B(s,\hat{\alpha},\beta) \leqslant W_B(s,\hat{\alpha},\hat{\beta})$ for every β, and Lemma 2 thus yields $W_A(s,\hat{\alpha},\hat{\beta}) \leqslant W_A(s,\hat{\alpha},\beta)$ for every β. Combining the inequalities, we have

$$\forall \alpha,\beta\colon\quad W_A(s,\alpha,\hat{\beta}) \leqslant W_A(s,\hat{\alpha},\hat{\beta}) \leqslant W_A(s,\hat{\alpha},\beta),$$

from which it clearly follows that

$$\min_\beta \max_\alpha W_A(s,\alpha,\beta) \leqslant W_A(s,\hat{\alpha},\hat{\beta}) \leqslant \max_\alpha \min_\beta W_A(s,\alpha,\beta). \tag{12}$$

Since $\max_x \min_y f(x,y) \leqslant \min_y \max_x f(x,y)$ is always true for any real-valued function f, it follows from (12) that

$$\max_\alpha \min_\beta W_A(s,\alpha,\beta) = \min_\beta \max_\alpha W_A(s,\alpha,\beta) = W_A(s,\hat{\alpha},\hat{\beta}) = W_A^v(s).$$

In order to establish that every minimax-stationary strategy $\hat{\alpha}$ for Player A satisfies $\hat{\alpha} \in \arg\min_\alpha \max_\beta W_B(s,\alpha,\beta)$, suppose to the contrary that for some minimax-stationary $\hat{\alpha}$ there is an α' that satisfies $\max_\beta W_B(s,\alpha',\beta) < \max_\beta W_B(s,\hat{\alpha},\beta)$. Since $\hat{\beta}$ is a best response to $\hat{\alpha}$ we have that $\max_\beta W_B(s,\hat{\alpha},\beta) = W_B(s,\hat{\alpha},\hat{\beta})$. Hence $\max_\beta W_B(s,\alpha',\beta) < W_B(s,\hat{\alpha},\hat{\beta})$, and so $\min_\alpha \max_\beta W_B(s,\alpha,\beta) < W_B(s,\hat{\alpha},\hat{\beta})$, which is a contradiction.

In order to establish that every minimax-stationary strategy $\hat{\alpha}$ for Player A satisfies $\hat{\alpha} \in \arg\max_\alpha \min_\beta W_A(s,\alpha,\beta)$, suppose to the contrary that for some minimax-stationary $\hat{\alpha}$ there is an α' that satisfies $\min_\beta W_A(s,\alpha',\beta) > \min_\beta W_A(s,\hat{\alpha},\beta)$. Since we have shown above that $\min_\beta W_A(s,\hat{\alpha},\beta) \geqslant W_A(s,\hat{\alpha},\hat{\beta})$ and that $W_A(s,\hat{\alpha},\hat{\beta}) = \max_\alpha \min_\beta W_A(s,\alpha,\beta)$, we clearly have a contradiction. □

The BMG Minimax Theorem establishes that by simply adopting a minimax-stationary behavior strategy, a player can guarantee himself a probability of at least $W_A^v(s)$ or $W_B^v(s)$ of winning the match, from whatever state s the match begins in. The following corollary establishes that after any history of play in the match (for example, even if the players have *not* so far been playing minimax-stationary strategies), a player can still guarantee himself a probability of at least $W_A^v(s)$ or $W_B^v(s)$ of winning the match, if the current state is s, by adopting a minimax-stationary strategy for the continuation game. The corollary, and the Uniqueness Theorem in Section 6, will require some notation:

Let $\overline{W}_A(h_t,\alpha,\beta)$ denote the probability that Player A will eventually win the match if the history is h_t and the players are following the behavior strategies α and β. (We do not require that h_t be consistent with α and β – *i.e.*, that the strategies α and β could yield h_t.) Let $h_t = (s_0, a_0, b_0, \ldots, a_{t-1}, b_{t-1}, s_t)$ and $h'_u = (s'_0, a'_0, b'_0, \ldots, a'_{u-1}, b'_{u-1}, s'_u)$ be histories such that $s_t = s'_0$. Then we write $h_t + h'_u$ for the history

$$\left(s_0, a_0, b_0, \ldots, a_{t-1}, b_{t-1}, s_t, a'_0, b'_0, \ldots, a'_{u-1}, b'_{u-1}, s'_u\right).$$

Corollary. *Let $h_t = (s_0, a_0, b_0, \ldots, a_{t-1}, b_{t-1}, s_t)$ be an arbitrary history. If Player A plays a minimax-stationary strategy after h_t, then his probability of winning the match is at least as great as his value when the state is s_t —* i.e., *if $\hat{\alpha}$ is a minimax-stationary strategy for Player A, then $\overline{W}_A(h_t,\hat{\alpha},\beta) \geqslant W_A^v(s_t)$ for every (not necessarily stationary) behavior strategy β for Player B.*

Proof. By the BMG Minimax Theorem we know the corollary is true for $t = 0$, *i.e.*, for every history h_0. Thus, let $t > 0$; let $h_t = (s_0, a_0, b_0, \ldots, a_{t-1}, b_{t-1}, s_t)$; let $\hat{\alpha}$ be a minimax-stationary strategy for Player A; and let β be a behavior strategy β for Player B that is consistent with h_t. Define a strategy β' as follows: $\beta'(h'_u) = \beta(h_t + h'_u)$ if $s'_0 = s_t$; and $\beta'(h'_u)$ is arbitrary for histories h'_u in which $s'_0 \neq s_t$. Since $\hat{\alpha}$ is a minimax-stationary strategy, we have $\overline{W}_A(h_t,\hat{\alpha},\beta) = W_A(s_t,\hat{\alpha},\beta') \geqq W_A^v(s_t)$, where the inequality follows from the BMG Minimax Theorem. □

6. The Uniqueness Theorem

We now establish a converse to the previous results: not only is it an equilibrium for both players to play minimax point-game strategies in every state, but this is the *only* equilibrium of the BMG if the point games themselves have unique equilibria.

Uniqueness Theorem. *Suppose that a binary Markov game satisfies the Monotonicity Condition and that, for each non-terminal state s, the associated point game G_s has a unique Nash equilibrium $(\sigma_A^*(s), \sigma_B^*(s))$ and $0 < v_A(s) < 1$. If (α^*, β^*) is a Nash equilibrium of the binary Markov game, then for each t and each equilibrium-path history $h_t \in H_t$, we have $\alpha^*(h_t) = \sigma_A^*(s_t)$ and $\beta^*(h_t) = \sigma_B^*(s_t)$.*

Corollary. *If each point game has a unique Nash equilibrium in which each player's strategy is completely mixed (i.e., every one of his pure strategies has positive mixture probability), then the binary Markov game has a unique Nash equilibrium.*

Proof of the Uniqueness Theorem. Let (α^*, β^*) be a Nash equilibrium of the binary Markov game. By the BMG Minimax Theorem, if the initial state is s, then every Nash equilibrium has the same value $W_A^v(s)$. The proof proceeds in three steps.

Step 1. We first establish that for every history h_t which is consistent with (α^*, β^*) (*i.e.*, which is on the (α^*, β^*)-equilibrium path), each player's probability of winning the match is $W_i^v(s_t)$, the value to him of the BMG if it were to begin in state s_t. Suppose to the contrary that for some history consistent with (α^*, β^*), one of the players does not obtain his value. Let t the first time at which there is an (α^*, β^*)-consistent history $h_t = (s_0, a_0, b_0, \ldots, a_{t-1}, b_{t-1}, s_t)$ such that $\overline{W}_i(h_t, \alpha^*, \beta^*) \neq W_i^v(s)$ for one of the players. By Lemma 2, we have $W_A^v(s) + W_B^v(s) = 1$, hence either $\overline{W}_A(h_t, \alpha^*, \beta^*) < W_A^v(s)$, or $\overline{W}_B(h_t, \alpha^*, \beta^*) < W_B^v(s)$, or both. Suppose $\overline{W}_A(h_t, \alpha^*, \beta^*) < W_A(s_t, \hat{\alpha}, \hat{\beta})$, and consider the strategy α' which coincides with α^* except that, for history h_t and all of its continuations, it plays a minimax strategy of the current state's point game. This deviation from α^* to α' is improving for Player A, *i.e.*, $W_A(s_0, \alpha', \beta^*) > W_A(s_0, \alpha^*, \beta^*)$ where s_0 is the initial state. This contradicts that (α^*, β^*) is a Nash equilibrium.

Step 2. We now establish that for every history h_t which is consistent with (α^*, β^*), Player A's probability of winning the current (period-t) point given (α^*, β^*) is $v_A(s_t)$. To lighten notation, we write s for the current state, s_t. We have

$$\overline{W}_A\big(h_t, \alpha^*, \beta^*\big) = \sum_{a\in A(s)} \sum_{b\in B(s)} \alpha^*(a|h_t)\beta^*(b|h_t)\pi_{sA}(a,b)\overline{W}_A\big(h_t + (a,b,s_+), \alpha^*, \beta^*\big)$$
$$+ \sum_{a\in A(s)} \sum_{b\in B(s)} \alpha^*(a|h_t)\beta^*(b|h_t)\pi_{sB}(a,b)\overline{W}_A\big(h_t + (a,b,s_-), \alpha^*, \beta^*\big).$$

By Step 1, since h_t is consistent with (α^*, β^*), we have $\overline{W}_A(h_t, \alpha^*, \beta^*) = W_A^v(s)$. Also by Step 1, if $h_t + (a,b,s_+)$ is consistent with (α^*, β^*), *i.e.*, if $\alpha^*(a|h_t)\beta^*(b|h_t)\pi_{sA}(a,b) > 0$, then $\overline{W}_A(h_t + (a,b,s_+), \alpha^*, \beta^*) = W_A^v(s_+)$ and similarly if $h_t + (a,b,s_-)$ is consistent with (α^*, β^*), then $\overline{W}_A(h_t + (a,b,s_-), \alpha^*, \beta^*) = W_A^v(s_-)$. Hence we can write

$$W_A^v(s) = W_A^v(s_+) \sum_{a\in A(s)} \sum_{b\in B(s)} \alpha^*(a|h_t)\beta^*(b|h_t)\pi_{sA}(a,b) + W_A^v(s_-) \sum_{a\in A(s)} \sum_{b\in B(s)} \alpha^*(a|h_t)\beta^*(b|h_t)\pi_{sB}(a,b). \quad (13)$$

Eq. (13), together with the inequality $W_A^v(s_+) > W_A^v(s_-)$ (the Monotonicity Condition) and the equation

$$W_A^v(s) = v_A(s)W_A^v(s_+) + v_B(s)W_A^v(s_-)$$

imply that

$$\sum_{a\in A(s)} \sum_{b\in B(s)} \alpha^*(a|h_t)\beta^*(b|h_t)\pi_{sA}(a,b) = v_A(s).$$

We have shown that if (α^*, β^*) is a Nash equilibrium, then for every history consistent with (α^*, β^*), the transitions from one state to the next must occur according to the value of the associated state. This completes Step 2.

Step 3. Continuing to write s for the state s_t attained via history h_t, we now establish that for every (α^*, β^*)-consistent history $h_t \in H_t$, we must have $\alpha^*(h_t) = \sigma_A^*(s)$ and $\beta^*(h_t) = \sigma_B^*(s)$, where $(\sigma_A^*(s), \sigma_B^*(s))$ is the unique Nash equilibrium of G_s. Suppose there is an equilibrium path history h_t such that $(\alpha^*(h_t), \beta^*(h_t)) \neq (\sigma_A^*(s), \sigma_B^*(s))$. Then Player A, say, has a mixed strategy $\sigma'_A \in \triangle A(s)$ such that

$$\sum_{a\in A(s)} \sum_{b\in B(s)} \sigma'_A(a)\beta^*(b|h_t)\pi_{sA}(a,b) > v_A(s).$$

Consider the strategy α' which is the same as α^* except at history h_t and its continuations, where we have $\alpha'(h_t) = \sigma'_A$ and, for all continuations of h_t – *i.e.*, for histories of the form $h_u = h_t + (s, a_t, b_t, \ldots, a_{u-1}, b_{u-1}, s_u)$ – α' satisfies $\alpha'(h_u) = \sigma_A^*(s_u)$. Then we have

$$\begin{aligned}\overline{W}_A\big(h_t, \alpha', \beta^*\big) &\geqslant W_A^v(s_+) \sum_{a\in A(s)} \sum_{b\in B(s)} \sigma'_A(a)\beta^*(b|h_t)\pi_{sA}(a,b) + W_A^v(s_-) \sum_{a\in A(s)} \sum_{b\in B(s)} \sigma'_A(a)\beta^*(b|h_t)\pi_{sB}(a,b) \\ &> v_A(s)W_A^v(s_+) + v_B(s)W_A^v(s_-) \\ &= \overline{W}_A\big(h_t, \alpha^*, \beta^*\big).\end{aligned}$$

The weak inequality follows from the Corollary to the Minimax Theorem; the strict inequality is implied by the Monotonicity Condition and the fact that Player A wins the point at history h_t with probability greater than $v_A(s)$; and the equation follows from Step 2.

We have shown that $\overline{W}_A(h_t, \alpha', \beta^*) > \overline{W}_A(h_t, \alpha^*, \beta^*)$. This implies that $W_A(s_0, \alpha', \beta^*) > W_A(s_0, \alpha^*, \beta^*)$, since Player A's probability of winning the match is the same when following α' as when following α^* if history h_t is not reached, but is greater when following α' if h_t is reached (which occurs with positive probability, since it is an equilibrium path history). This contradicts that (α^*, β^*) is a Nash equilibrium. □

7. Examples

The first example is a parametric class of binary Markov games, in which the Monotonicity Condition holds for some parameter values and not for others. When the Monotonicity Condition fails in these games, it is not a Nash equilibrium of the binary Markov game (the match) for each player to always (*i.e.*, in each state) play the minimax equilibrium of the state's point game. This class of games demonstrates that the Monotonicity Condition cannot be dispensed with in either the Equilibrium Theorem or the Minimax Theorem.

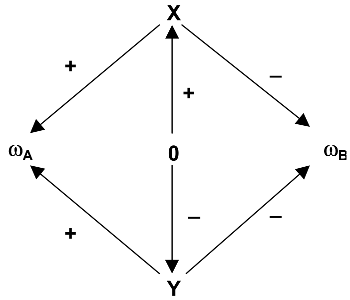

Fig. 3.

Example 4. Consider the binary scoring rule illustrated in Fig. 3, in which the state space is $S = \{0, X, Y, \omega_A, \omega_B\}$, the initial state is 0, and the two transition functions are given by $0_+ = X$, $0_- = Y$, $X_+ = Y_+ = \omega_A$, and $X_- = Y_- = \omega_B$. For each state s the matrix G_s gives Player A's payoff function π_{sA}, where x and y are parameters satisfying $0 < x, y \leqslant \frac{1}{2}$:

	L	R	
T	1	0	0
M	3/4	1/2	0.8
B	0	1	0.2
	0.4	0.6	

G_0

	L	R	
T	0	$2x$	0.5
B	$2x$	0	0.5
	0.5	0.5	

$G_X(x)$

	L	R	
T	0	$2y$	0.5
B	$2y$	0	0.5
	0.5	0.5	

$G_Y(y)$

Player A's (B's) minimax mixture is given to the right (on the bottom) of the payoff matrix. We have $v_A(0) = 0.6$, $v_A(X) = x$, and $v_A(Y) = y$.

For each state s, writing as before $W_A^v(s)$ for the probability that Player A wins the match when the transition at each state occurs according to the point game's value, we have that $W_A^v(0) = 0.6x + 0.4y$, $W_A^v(X) = x$, and $W_A^v(Y) = y$. The monotonicity condition holds if $W_A^v(X_+) > W_A^v(X_-)$, $W_A^v(Y_+) > W_A^v(Y_-)$, and $W_A^v(0_+) > W_A^v(0_-)$ (and similarly for Player B). The first two inequalities always hold since $1 = W_A^v(\omega_A) > W_A^v(\omega_B) = 0$. The third holds if and only if $x > y$. If $x > y$, then by the Equilibrium Theorem it is a Nash equilibrium for each player to play minimax (and Nash) at each state, *i.e.*, (α^*, β^*) is a Nash equilibrium where $\alpha^*(0) = (0, 0.8, 0.2)$, $\alpha^*(X) = \alpha^*(Y) = (0.5, 0.5)$, $\beta^*(0) = (0.8, 0.2)$ and $\beta^*(X) = \beta^*(Y) = (0.5, 0.5)$. Since each point game has a unique Nash equilibrium, by the Uniqueness Theorem (α^*, β^*) is the unique Nash equilibrium of the BMG.

If $y > x$, and so the monotonicity condition fails, then (α^*, β^*) is not a Nash equilibrium: Consider the strategy α' which is the same as α^*, except that at the initial state Player A chooses T for sure, *i.e.*, $\alpha'(0) = (1, 0, 0)$. Employing α' against β^* Player A wins the match with probability $0.6y + 0.4x$, whereas if he employs α^* he wins with only probability $0.6x + 0.4y$. Hence α^* is not a best response to β^*. (When $y > x$ then $\hat{\alpha}(0) = (\frac{1}{2}, 0, \frac{1}{2})$, $\hat{\alpha}(X) = \hat{\alpha}(Y) = (0.5, 0.5)$, and $\hat{\beta}(0) = \hat{\beta}(X) = \hat{\beta}(Y) = (0.5, 0.5)$ is the unique Nash equilibrium.)

The next example shows that if, at some state, the associated point game has more than two outcomes, then equilibrium play in the point game may depend upon the values of the point games at other states. In contrast, in binary Markov games, so long as the Monotonicity Condition is satisfied, equilibrium play at every point game is independent of the values of the point games at other states.

Example 5. Consider the family of games illustrated in Fig. 4, for values of x satisfying $0 < x \leqslant \frac{1}{2}$. The state space is $S = \{0, D, \omega_A, \omega_B\}$, where 0 is the initial state. At the initial state the point game has three possible outcomes: Player A

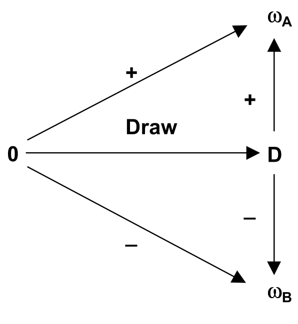

Fig. 4.

		Player B	
		L	R
Player A	L	π_{sA}(L,L)	π_{sA}(L,R)
	R	π_{sA}(R,L)	π_{sA}(R,R)

Fig. 5. A typical point game in tennis. $\pi_{sA}(a, b)$ is the probability that Player A will win the point if Player A chooses action a and Player B chooses action b.

wins the point, Player B wins the point, or neither player wins the point (*i.e.*, the outcome is a "draw"). The transition function is given by $0_+ = D_+ = \omega_A$, $0_- = D_- = \omega_B$, and the match transits to state D if the outcome at the initial state is a draw. The first player to win a point wins the match. This is not a binary Markov game: the point game at the initial state has more than two outcomes.

The entries in the matrices G_0 and $G_D(x)$ below give the players' probabilities of winning the point. Note that neither player wins the point at the initial state if the action profile is (B, R):

$$\begin{array}{c|c|c|}
 & \text{L} & \text{R} \\ \hline
\text{T} & 1,0 & 0,1 \\ \hline
\text{B} & 0,1 & 0,0 \\ \hline
\end{array}
\qquad
\begin{array}{c|c|c|c}
 & \text{L} & \text{R} & \\ \hline
\text{T} & 2x, 1-2x & 0, 1 & 0.5 \\ \hline
\text{B} & 0, 1 & 2x, 1-2x & 0.5 \\ \hline
 & 0.5 & 0.5 &
\end{array}$$

$$G_0 \qquad\qquad G_D(x)$$

The point game at state D has value x to the Row player and $1 - x$ to the Column player. Replacing the zero probabilities in the (B, R) cell of G_0 with x and $1 - x$ makes it easy to see that the overall game has a unique Nash equilibrium (α^*, β^*) in which $\alpha^*(0) = \beta^*(0) = (\frac{x}{1+x}, \frac{1}{1+x})$ and $\alpha^*(D) = \beta^*(D) = (0.5, 0.5)$. In a binary Markov game, Nash equilibrium play at the initial state depends only on the probabilities of the players winning the initial point, but in this game Nash equilibrium play at the initial state also depends upon the players' values of the point game at state D.

8. An application: The game of tennis[13]

We describe here a model of a tennis match as a binary Markov game,[14] and we show that the game satisfies the Monotonicity Condition. Therefore all of our results apply: it is a Nash equilibrium in our model of a tennis match for each player to play a minimax strategy at each point; moreover, it is a minimax behavior strategy in the match for a player to play in this way; and, since each point game in our tennis model has a unique equilibrium, the Uniqueness Theorem guarantees that the *only* equilibrium in the match — *i.e.*, in the binary Markov game — is for each player to play his unique minimax strategy in every point game that arises in the match.

Each point in a tennis match is begun by one of the players placing the ball in play, or "serving." Our model of a tennis match focuses on the decisions by the server and the receiver on the serve. We assume that the two actions available to the server are to serve either to the receiver's Left (L) or to the receiver's Right (R). Simultaneously with the server's decision, the receiver is assumed to guess whether the serve will be delivered to his Left or to his Right. Thus, for every state s (we will describe the states shortly), we have $A(s) = B(s) = \{L, R\}$.

After the players have made their Left-or-Right choices for the serve, many subsequent strokes may be required (in an actual tennis match) to determine which player is the winner of the current point. We leave this after-the-serve part of the point unmodeled, and instead adopt a reduced-form model of the point, as depicted in Fig. 5: each player's payoffs in the four cells of the game matrix are the respective probabilities that he will *ultimately* win the point at hand, conditional on the Left-or-Right choices each of the players has made on the serve. Player A's payoffs, $\pi_{sA}(a, b)$, are shown in Fig. 5; Player B's payoff in each cell is $1 - \pi_{sA}(a, b)$. For each state s in which Player i is the server, we naturally assume that $\pi_{si}(L, L) < \pi_{si}(L, R)$, $\pi_{si}(R, R) < \pi_{si}(R, L)$, $\pi_{si}(L, L) < \pi_{si}(R, L)$, and $\pi_{si}(R, R) < \pi_{si}(L, R)$. Hence, in every point game G_s, whichever player is serving, each player has a unique minimax strategy, which is strictly mixed, and the value of the point game to each player is strictly between 0 and 1.

The scoring of a tennis match is structured as follows. The match consists of a sequence of "sets": when a set is completed (*i.e.*, when one of the players has won the set), then the next set begins. The first player to win three sets is

---

[13] This section of the paper forms the theoretical foundation for Walker and Wooders (2001).

[14] The binary Markov game model of tennis assumes that the probabilities of A or B winning a given point depend only on the current score. This score, or state variable, need not correspond precisely to the score as it is usually understood, but might also include such factors as whether the game score is "even" or "odd" (*i.e.*, a deuce-court or an ad-court point), which player is serving, which end of the court the server is playing from (if sun or wind is a factor), and even whether it is early in the match or late by distinguishing first-set points from second-set points, etc. But it assumes that *only* this current score matters, which would not be the case if, for example, the actions a player chose earlier in the match determine how much energy he has in reserve for later points – a plausible factor in an actual tennis match. This is an interesting empirical question, and the results in this paper form the theoretical foundation for such an empirical analysis.

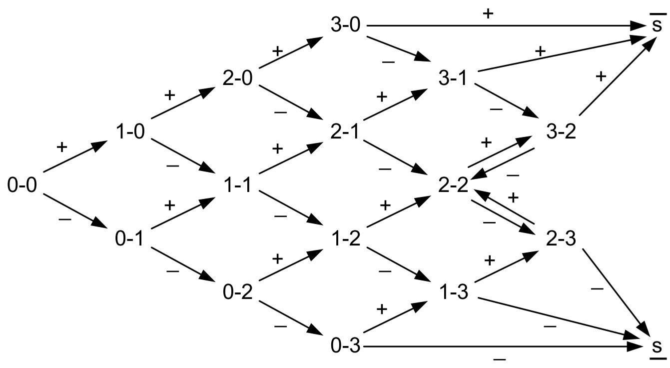

Fig. 6. The scoring rule for a "game" in tennis.

declared the winner of the match. (In some matches it is the first player to win two sets.) Each set is in turn composed of a sequence of "games": a set is won by the first player to win at least six games and to simultaneously have won two more games than his opponent. And each game is composed of a sequence of points: a game is won by the first player to win at least four points and to have won at least two more points than his opponent.[15] The score at any juncture of a match is thus given by specifying (a) the set score, *i.e.*, how many sets each player has won; (b) the game score, *i.e.*, how many games each player has won in the current set, or, if each has won at least four games, then simply whether the players have won the same number of games, or whether Player A or Player B is ahead by one game; and (c) the point score, *i.e.*, how many points each player has won in the current game, or, if each has won at least two points, then simply whether the players have won the same number of points, or whether Player A or Player B is ahead by one point. This latter component of the score, the point score in a "game," is described in Example 3 in Section 1, and is depicted in Fig. 6. (Fig. 6 is identical to Fig. 2, except that here in Fig. 6 the "game" in tennis is not the entire match, or binary Markov game, so that the game's terminal states are not typically the terminal states ω_A and ω_B of the match. We instead use the notation $\bar{s}$ for the state the match will be in if Player A wins the current game, and $\underline{s}$ for the state the match will be in if Player A loses the game.)

The set of all possible scores, as described in (a), (b), and (c) above, is finite and comprises the set S of states in our binary Markov game model of a tennis match. The transition functions are defined in the obvious way: for any state (*i.e.*, score) s, the states s_+ and s_- are the scores reached if Player A wins or loses the current point.

Our objective is to verify that the match satisfies the Monotonicity Condition, *i.e.*, to show that a player's probability of winning the match is always greater if he wins the current point than if he loses the point, if the state transitions occur according to the values of the point games. We begin by assuming (for the moment) that for every possible set score and game score, a player's probability of winning the match is always greater if he wins the current game than if he loses it, and we show that the Monotonicity Condition then follows.

The progression of scores in the current game is depicted in Fig. 6. Each node in the figure corresponds to one of the possible scores in the game[16] and thus to one of the possible states in the match (given the game score and the set score). The state transitions in Fig. 6 are assumed to occur according to the values of the point games at the various nodes, or states — namely, $v_A(s)$ is the probability of moving to s_+ and $v_B(s)$ is the probability of moving to s_-. Note that $v_A(s) + v_B(s) = 1$. To simplify notation, we write v_s for the probability of moving to state s_+ and $W(s)$ for the probability Player A will win the match if the match is in state s; and we denote the various scores, or states s, by 00, 10, 12, *etc.*, in place of 0–0, 1–0, 1–2. For each state s, then, we have

$$0 < v_s < 1 \tag{14}$$

and

$$W(s) = v_s W(s_+) + (1 - v_s) W(s_-), \tag{15}$$

and we are (for the moment) assuming that

$$W(\underline{s}) < W(\bar{s}). \tag{16}$$

The states $\bar{s}$ and $\underline{s}$, while not absorbing states in the match, are absorbing states in Fig. 6. We refer to the rightmost column in Fig. 6, which contains $\bar{s}$ and $\underline{s}$, as the "absorbing column." It will be useful to write $s < t$ when two non-absorbing

[15] Which of the players is the server during the first game of the match is determined by lot, and then the players reverse roles (server and receiver) at the beginning of each new game.

[16] It is a convention in tennis that a player's score is said to be 15 (instead of 1) when he has won one point; and to be 30, and then 40, when he has won two or three points. We express the score instead simply in terms of the number of points each player has won.

states s and t lie in the same column of Fig. 6 (*i.e.*, if the two states can be reached via the same number of points) and if s lies below t in the column (*i.e.*, if Player A has won more points in state t than in state s). Notice that

$$s < t \quad \Rightarrow \quad \big[(s_+ < t_+ \text{ or } t_+ = \bar{s}) \text{ and } (s_- < t_- \text{ or } s_- = \underline{s})\big].$$

It is clear from (14) and (15) that the Monotonicity Condition is equivalent to the following:

$$s < t \quad \Rightarrow \quad W(s) < W(t). \tag{17}$$

Hence, we establish (17).

First, we verify that (17) is satisfied for the states $s = 23$ and $t = 32$ — *i.e.*, for the two states in the rightmost non-absorbing column of Fig. 6. From (15), we have the following system of three equations in the winning probabilities $W(\cdot)$:

$$W(32) = v_{32}W(\bar{s}) + (1 - v_{32})W(22),$$

$$W(22) = v_{22}W(32) + (1 - v_{22})W(23),$$

$$W(23) = v_{23}W(22) + (1 - v_{23})W(\underline{s}).$$

It clearly follows from (14), (16), and these three equations that

$$W(\underline{s}) < W(23) < W(22) < W(32) < W(\bar{s}),$$

which verifies (17) for the rightmost non-absorbing column in Fig. 6.

Now we show that if (17) holds for one of the columns in Fig. 6, then it holds as well for the column immediately to the left. Consider two adjacent columns, and let s and t be two states that lie in the left one, with $s < t$; we must show that $W(s) < W(t)$. We have

$$W(t) = v_t W(t_+) + (1 - v_t)W(t_-), \quad \text{and} \quad W(s) = v_s W(s_+) + (1 - v_s)W(s_-). \tag{18}$$

There are three cases to consider: (i) $t_+ \neq \bar{s}$ and $s_- \neq \underline{s}$, (ii) $t_+ = \bar{s}$, and (iii) $s_- = \underline{s}$. In the first case, s_+, s_-, t_+ and t_- all lie in the column immediately to the right of s and t, and we have assumed that (17) holds in that column. We have $s_- < s_+ \leqslant t_- < t_+$, and therefore (14) and (18) yield $W(s) < W(t)$. In case (ii) we have $W(s_+) \leqslant W(t_+)$ and $s_- < t_-$, so that (14) and (18) again yield $W(s) < W(t)$. Similarly, in case (iii) we have $W(s_-) \leqslant W(t_-)$ and $s_+ < t_+$, so that (14) and (18) yield $W(s) < W(t)$. Thus, we have established that indeed, if (17) holds in a non-absorbing column of Fig. 6, then it also holds for the next column to the left. Since we have also established that (17) holds in the rightmost non-absorbing column, it follows that it holds for every non-absorbing column of Fig. 6, and therefore the match satisfies the Monotonicity Condition — under the assumption that $W(\underline{s}) < W(\bar{s})$.

Now assume that each player always obtains a greater probability of winning the match by winning the current set than by losing it. Since the winning of games now determines who wins the set in exactly the same way that the winning of points determined who wins the game (except that one must win six games to win the set, but only four points to win the game), the same argument as above clearly establishes that a player will obtain a greater probability of winning the match if he wins the current game than if he loses it — *i.e.*, that $W(\underline{s}) < W(\bar{s})$. Therefore the match satisfies the Monotonicity Condition if each player's probability of winning the match is always greater after winning a set than after losing it. And finally, it is easy to see that a player's probability of winning the match is indeed greater if he wins any given set than if he loses it, which (as we have just described) ensures that $W(\underline{s}) < W(\bar{s})$, and therefore the Monotonicity Condition is indeed satisfied for the match.

References

Blackwell, D., Ferguson, T.S., 1968. The big match. Ann. of Math. Statist. 39, 159–163.

Everett, H., 1957. Recursive games. In: Dresher, M., Tucker, A.W., Wolfe, P. (Eds.), Contributions to the Theory of Games. In: Ann. of Math. Stud., vol. III. Princeton University Press, Princeton, NJ, pp. 47–78.

Flesch, J., Thuijsman, F., Vrieze, O.J., 1996. Recursive repeated games with absorbing states. Math. Operations Res. 21, 1016–1022.

Mertens, J.-F., 2002. Stochastic games. In: Aumann, R., Hart, S. (Eds.), Handbook of Game Theory, vol. 3. Elsevier, Amsterdam, pp. 1810–1832.

Mertens, J.-F., Neyman, A., 1981. Stochastic games. Int. J. Game Theory 10, 53–66.

Milnor, J., Shapley, L., 1957. Games of survival. In: Dresher, M., Tucker, A.W., Wolfe, P. (Eds.), Contributions to the Theory of Games. In: Ann. Math. Stud., vol. III. Princeton University Press, Princeton, NJ, pp. 15–45.

Neyman, A., Sorin, S., 2003. Stochastic Games and Applications. Kluwer, Dordrecht, Boston.

Shapley, L., 1953. Stochastic games. Proc. Natl. Acad. Sci. USA 39, 1095–1100.

Tijs, S., Vrieze, O.J., 1986. On the existence of easy initial states for undiscounted stochastic games. Math. Operations Res. 11, 506–513.

Thuijsman, F., Vrieze, O.J., 1991. Easy initial states for undiscounted stochastic games. In: Raghavan, T.E.S., Parthasarathy, T., Ferguson, T.S., Vrieze, O.J. (Eds.), Stochastic Games and Related Topics. Kluwer Academic, Norwell, MA, Dordrecht, pp. 85–99.

Vieille, N., 2000a. Two-player stochastic games I: A reduction. Israel J. Math. 119, 55–91.

Vieille, N., 2000b. Two-player stochastic games II: The case of recursive games. Israel J. Math. 119, 93–126.

Vieille, N., 2002. Stochastic games: Recent results. In: Aumann, R., Hart, S. (Eds.), Handbook of Game Theory, vol. 3. Elsevier, Amsterdam, pp. 1833–1850.

Walker, M., Wooders, J., 2001. Minimax play at Wimbledon. Amer. Econ. Rev. 91, 1521–1538.

Wooders, J., Shachat, J., 2001. On the irrelevance of risk attitudes in repeated two-outcome games. Games Econ. Behav. 34, 342–363.