The Annals of Mathematical Statistics
1968, Vol. 39, No. 1, 159–163

THE BIG MATCH[1]

BY DAVID BLACKWELL AND T. S. FERGUSON

University of California, Berkeley and Los Angeles

1. Introduction. Extending Shapley's work on stochastic games [2], Gillette [1] has studied the following situation. We are given three non-empty finite sets S, I, J, a real-valued function a defined for all triples (s, i, j), $s \varepsilon S$, $i \varepsilon I$, $j \varepsilon J$, and a function p which associates with each triple (s, i, j) a probability distribution $p(\cdot \mid s, i, j)$ on S. These five quantities S, I, J, a, p define a two-person zero-sum game, played as follows. We start with some initial state $s \varepsilon S$, known to both players. Player 1 chooses $i \varepsilon I$ and, simultaneously, player 2 chooses $j \varepsilon J$. Player 1 is then awarded $a(s, i, j)$ points, and the game moves to state s' selected according to $p(\cdot \mid s, i, j)$. The new state s' is announced to both players, who then choose i', j', giving player 1 $a(s', i', j')$ points, and causing the game to move to the state s'' selected according to $p(\cdot \mid s', i', j')$, etc. The payoff to player 1 from the infinite sequence of choices is

$$\limsup_{n\to\infty} (a_1 + \cdots + a_n)/n,$$

where a_m is the point score 1 obtained on the mth round. Whether such games, the finite stochastic games, always have a value is not known. In this paper we consider one interesting example of Gillette, the big match, and show that it does have a value.

The big match is played as follows. Every day player 2 chooses a number, 0 or 1, and player 1 tries to predict 2's choice, winning a point if he is correct. This continues as long as player 1 predicts 0. But if he ever predicts 1, all future choices for both players are required to be the same as that day's choices: if player 1 is correct on that day, he wins a point every day thereafter; if he is wrong on that day, he wins zero every day thereafter. The payoff to 1 is

$$\limsup_{n\to\infty} (a_1 + \cdots + a_n)/n,$$

where a_m is the number of points he wins on the mth day. The big match is the finite stochastic game with

$$S = \{0, 1, 2\}, \qquad I = J = \{0, 1\},$$

$$a(2, i, j) = \delta_{ij}, \qquad a(s, i, j) = s \quad \text{for} \quad s = 0, 1,$$

$$p(2, 0, j) = \delta(2), \qquad p(2, 1, j) = \delta(j),$$

$$p(s, i, j) = \delta(s) \quad \text{for} \quad s = 0, 1, \quad \text{and initial position } s = 2.$$

Received 13 June 1967.

[1] Supported in part by the Office of Naval Research under Contract Nonr-222 (53), and in part by National Science Foundation Grant GP-5224.

2. Solution of the big match.

THEOREM 1. *The value of the big match is* $\frac{1}{2}$. *An optimal strategy for player 2 is to toss a fair coin every day. Player 1 has no optimal strategy, but for any non-negative integer N he can get*

$$V(N) = N/2(N+1)$$

by using strategy N, defined as follows: having observed player 2's first n choices $x_1, \cdots, x_n$, $n \geqq 0$, *calculate the excess* k_n *of* 0's *over* 1's *among* $x_1, \cdots, x_n$, *and predict* 1 *with probability* $p(k_n + N)$, *where* $p(m) = 1/(m+1)^2$.

PROOF. Clearly if player 2 tosses a fair coin every day, player 1's expected payoff is $\frac{1}{2}$, no matter what he does: if he ever predicts 1 his payoff is equally likely to be 0 or 1, and if he predicts 0 forever, the strong law of large numbers gives him payoff $\frac{1}{2}$ with probability 1.

Next, notice that strategy N predicts 1 with certainty whenever the excess is $-N$, i.e. whenever N more 1's than 0's have occurred. Suppose we have verified that strategy N produces expected payoff at least $V(N)$ against every sequence of 0's and 1's which eventually achieves an excess of $-N$, and consider any sequence $\omega = (x_1, x_2, \cdots)$ which never achieves an excess as low as $-N$. We show that strategy N yields at least $V(N)$ against ω. Denote the number of observations after which player 1 first predicts 1 by t (so $t = \infty$ if he never predicts 1) Define

$$\lambda(m) = P\{t < m \text{ and } x_{t+1} = 0\},$$

$$\mu(m) = P\{t < m \text{ and } x_{t+1} = 1\},$$

$$\lambda = \lim_{m\to\infty} \lambda(m), \qquad \mu = \lim_{m\to\infty} \mu(m).$$

Player 1's expected income is at least

$$\mu + (1 - \lambda - \mu)(\tfrac{1}{2}),$$

since the sequence ω, never having excess as low as $-N$, satisfies

$$(x_1 + \cdots + x_n)/n < \tfrac{1}{2} + N/2n, \quad \text{for all } n$$

so that 1's income when $t = \infty$, which is

$$\limsup_{n\to\infty} ((1 - x_1) + \cdots + (1 - x_n))/n,$$

is at least $\frac{1}{2}$ (this would be true even with lim sup replaced by lim inf).

To show that $\mu + (1 - \lambda - \mu)(\frac{1}{2}) \geqq V(N)$, consider player 1's income from strategy N against the following strategy for 2: choose $x_1, \cdots, x_m$, and toss a fair coin thereafter. With probability 1, the resulting sequence will eventually reach excess $-N$ so, by assumption, player 1's expected income is at least $V(N)$. But his expected income is exactly

$$\mu(m) + \tfrac{1}{2}(1 - \lambda(m) - \mu(m)),$$

which therefore is at least $V(N)$. Letting $m \to \infty$ completes the proof that it suffices to study sequences ω which eventually achieve an excess of $-N$.

Take such an $\omega = (x_1, x_2, \cdots)$, let player 1 use strategy N against it, and denote by t the number of observations after which he first predicts 1. Define

$$E(m) = \{t \geqq m, \quad \text{or} \quad t < m \quad \text{and} \quad x_{t+1} = 1\}, \quad m = 1, 2, \cdots.$$

We show, inductively on m, that

$$(1) \qquad P_N(E(m)) \geqq V(N), \quad \text{for all } N.$$

(a) $m = 1$. If $x_1 = 1$, $P_N(E(1)) = 1 > V(N) = N/2(N+1)$. If $x_1 = 0$, $P_N(E(1)) = P_N\{t \geqq 1\} = 1 - p(N) = N(N+2)/(N+1)^2 \geqq V(N)$.

(b) Suppose $P_N(E(m)) \leqq V(N)$, for all N. If $x_1 = 1$, $P_N(E(m+1)) = p(N) + [1 - p(N)]\, P_{N-1}(E(m)) \geqq p(N) + [1 - p(N)]V(N-1) = V(N)$, where $P_{N-1}(E(m)) \geqq V(N-1)$ by induction since using strategy N against $\omega = (1, x_2, x_3, \cdots)$ is equivalent to predicting 1 initially with probability $p(N)$ and, with probability $1 - p(N)$ predicting 0 initially and thereafter using strategy $N - 1$ against $\omega' = (x_2, x_3, \cdots)$. Similarly, if $x_1 = 0$,

$$P_N(E(m+1)) = [1 - p(N)]P_{N+1}(E(m)) \geqq [1 - p(N)]V(N+1) = V(N).$$

So (1) is proved. Since $t < \infty$ with probability 1, letting $m \rightarrow \infty$ in (1) yields

$$P_N\{x_{t+1} = 1\} \geqq V(N),$$

strategy N yields at least $V(N)$ against every sequence, and the value of the game is $\frac{1}{2}$.

To show that 1 has no optimal strategy, consider any strategy σ for him. If σ never predicts 1 with positive probability against $\omega^* = (1, 1, 1, \cdots)$, it wins 0 against ω^* so is not optimal. If not, say $m \geqq 0$ is the smallest initial number of 1's after which σ predicts 1 with positive probability, say ϵ. Player 2 can counter σ by choosing m initial 1's, then 0, and tossing a fair coin thereafter, giving 1 an expected income of $(1 - \epsilon)\frac{1}{2}$.

The above argument that 1 has no optimal strategy is due to Lester Dubins. We are indebted to him, David Freedman, and Volker Strassen for stimulating conversations about the big match.

3. Other near optimal strategies for player 1. Let $0 < \epsilon < 1$, and let $\{\alpha_n\}$, $n = 0, 1, 2, \cdots$, be a sequence of numbers satisfying the conditions

(a) $\alpha_n \geqq \alpha_{n+1}$ for $n = 0, 1, 2, \cdots$,

(b) $(1 - \epsilon)\alpha_n \leqq \alpha_{n+1}$ for $n = 0, 1, 2, \cdots$,

(c) $\sum_0^\infty \alpha_n = 1$.

(Automatically, $\alpha_n > 0$ for all n.) For a given $0 < \epsilon < 1$ and $\{\alpha_n\}$ satisfying (a) (b) and (c), we define a strategy ψ for player 1 by defining the distribution of t, the time at which player 1 first predicts 1, via the functions

$$\psi_n(x_1, \cdots, x_n) = P\{t = n + 1 \mid x_1, \cdots, x_n\} \qquad n = 0, 1, 2, \cdots.$$

Note that ψ_n represents the (unconditional) probability that $t = n + 1$, whereas in Theorem 1 we referred to the conditional probabilities, $P\{t = n + 1 \mid t > n, x_1, \cdots, x_n\}$.

The functions ψ_n are defined inductively by letting $\psi_0 = \epsilon\alpha_0$ and

$$\psi_n(x_1, \cdots, x_n) = \epsilon\alpha_{k_n}, \qquad \text{if } \textstyle\sum_0^n \epsilon\alpha_{k_j} \leqq 1$$
$$= 1 - \textstyle\sum_0^{n-1} \psi_j(x_1, \cdots, x_j), \qquad \text{otherwise,}$$

where k_n, as before, represents the excess of 0's over 1's among $x_1, \cdots, x_n$ ($k_n = \sum_1^n (1 - 2x_i)$), and where for $j < 0$, α_j is defined equal to α_0. We denote by $\mathcal{C}_\epsilon$ the class of strategies ψ formed in this manner for some sequence $\{\alpha_n\}$ satisfying (a) (b) and (c).

The main property of the class $\mathcal{C}_\epsilon$ is as follows:

THEOREM 2. *If* $\psi \in \mathcal{C}_\epsilon$, $\sum_0^\infty (1 - 2x_{n+1})\psi_n(x_1, \cdots, x_n) \leqq 3\epsilon$.

PROOF. If $x_1, x_2, \cdots$ is such that $\sum_0^\infty \epsilon\alpha_{k_n} \leqq 1$ (so that $\psi_n = \epsilon\alpha_{k_n}$), there is an integer M such that $\epsilon \sum_0^{M-1} \alpha_{k_n} \geqq \epsilon \sum_0^\infty \alpha_{k_n} - \epsilon$. If not, let M denote the smallest integer for which $\sum_0^M \epsilon\alpha_{k_n} > 1$ (so that $\psi_n = \epsilon\alpha_{k_n}$ for $n < M$). In either case,

$$\sum_0^\infty (1 - 2x_{n+1})\psi_n(x_1, \cdots, x_n) \leqq \epsilon\sum_0^{M-1}(1 - 2x_{n+1})\alpha_{k_n} + \epsilon.$$

Let $J = k_M$. We assume $J \geqq 0$. (For $J < 0$, a similar argument obviously works.) Let

$$I_1 = \{n\colon k_{n+1} > \max\nolimits_{0 \leqq j \leqq m} k_j, \quad k_n < J, \quad 0 \leqq n \leqq M - 1\},$$
$$I_2 = \{n\colon x_{n+1} = 0, \quad n \notin I_1, \quad 0 \leqq n \leqq M - 1\},$$
$$I_3 = \{n\colon x_{n+1} = 1, \quad 0 \leqq n \leqq M - 1\}.$$

There is a one-to-one correspondence between I_2 and I_3 : If $k_i \leqq J$, $i \in I_3$ is paired with the smallest $j \in I_2$ such that $j > i$ and $k_j = k_i - 1$. If $k_i > J$, $i \in I_3$ is paired with the largest $j \in I_2$ such that $j < i$ and $k_j = k_i - 1$. In this pairing, $\alpha_{k_i} \geqq (1 - \epsilon)\alpha_{k_j}$. Hence,

$$\sum_0^{M-1} (1 - 2x_{n+1})\alpha_{k_n} = \sum_{I_1} \alpha_{k_n} + \sum_{I_2} \alpha_{k_j} - \sum_{I_3} \alpha_{k_i} \leqq 1 + \epsilon \sum_{I_2} \alpha_{k_j} \leqq 2$$

completing the proof.

In the big match, the expected payoff to player 1 given $x_1, x_2, \cdots$ is, in terms of the functions $\{\psi_n\}$,

$$\sum_0^\infty x_{n+1}\psi_n(x_1, \cdots, x_n) + \left(1 - \sum_0^\infty \psi_n(x_1, \cdots, x_n)\right) \limsup (1 - \bar{x}_n)$$

where $\bar{x}_n$ represents $(x_1 + \cdots + x_n)/n$. Theorem 2 implies that for $\psi \in \mathcal{C}_\epsilon$ this expected payoff is at least

$$\tfrac{1}{2}\sum_0^\infty \psi_n(x_1, \cdots, x_n) - \tfrac{3}{2}\epsilon + \left(1 - \sum_0^\infty \psi_n(x_1, \cdots, x_n)\right) \limsup (1 - \bar{x}_n).$$

Hence, if $\limsup (1 - \bar{x}_n) \geqq \frac{1}{2}$, the expected payoff is at least $\frac{1}{2} - (\frac{3}{2})\epsilon$. On the other hand, if $\limsup (1 - \bar{x}_n) < \frac{1}{2}$, then k_n is negative infinitely often so that $\sum_0^\infty \psi_n(x_1, \cdots, x_n) = 1$ for any $\psi \in \mathcal{C}_\epsilon$. Thus, in any case, the expected payoff is at least $\frac{1}{2} - (\frac{3}{2})\epsilon$. In other words, every $\psi \in \mathcal{C}_\epsilon$ is $(\frac{3}{2})\epsilon$-optimal for the big match.

In closing, we mention a related problem, in which further restrictions are placed on the strategies available to the player. We require of player 2 that he choose a sequence $\omega = (x_1, x_2, \cdots)$ such that $\lim \bar{x}_n = \frac{1}{2}$, and we require of player 1 that he choose a distribution of t such that $P\{t < \infty \mid x_1, x_2, \cdots\} = 1$ for all ω available to player 2. The value of this modified game is still $\frac{1}{2}$, as may be seen from the following considerations. Player 2's optimal strategy for the big match is still available to him in the modified game, so the upper value is at most $\frac{1}{2}$. To see that the lower value is at least $\frac{1}{2}$, let player 1 play as follows. He chooses a small $\delta > 0$ and produces a private random sequence, $y_1, y_2, \cdots$ of independent variables with $P\{y_i = 1\} = \delta$ and $P\{y_i = 0\} = 1 - \delta$. He defines $z_n = \max(x_n, y_n)$ and uses any strategy in $\mathcal{C}_\epsilon$ (or the strategy in Theorem 1), pretending that 2's sequence is $z_1, z_2, \cdots$. Since $\bar{x}_n \to \frac{1}{2}$, $\bar{z}_n \to (1 + \delta)/2$, so that $t < \infty$ with probability one. Then, $P\{z_{t+1} = 1\} > \frac{1}{2} - (\frac{3}{2})\epsilon$. But since $P\{y_{t+1} = 1\} = \delta$, we have $P\{x_{t+1} = 1\} > \frac{1}{2} - (\frac{3}{2})\epsilon - \delta$, showing the lower value is at least $\frac{1}{2}$.

REFERENCES

[1] Gillette, Dean (1957). Stochastic games with zero stop probabilities. *Contributions to the Theory of Games* **3**. Princeton University Press.

[2] Shapley, L. S. (1953). Stochastic games. *Proc. Nat. Acad. Sci.* **39** 1095–1100.