The Big Match

Joshua Ruiter
Virginia Wesleyan University

December, 2025

Outline

The matching game

In the **matching game**, players A and B each hide 1 or 2 fingers, then simultaneously reveal.

- Match $\to +1$ point for A and -1 point for B
- Mismatch $\to -1$ for A and $+1$ for B

	1	2
1	$(1,-1)$	$(-1,1)$
2	$(-1,1)$	$(1,-1)$

	1	2
1	1	-1
2	-1	1

A chooses the row, B chooses the column.

Activity time! Find a neighbor, decide on roles A and B, and play 10 rounds.

Pure strategies

Definition

A **pure strategy** in a game is a deterministic strategy which involves always making the same deterministic decision in a given situation.

Example

In the matching game, the pure strategies are "Always play 1" and "Always play 2."

Mixed strategies

Definition

A **mixed strategy** is a strategy which involves randomizing among some set of pure strategies, with some kind of weighted probability.

Example

In the matching game, a mixed strategy is a strategy of the form

"Play 1 with probability p_1, and play 2 with probability $p_2 = 1 - p_1$"

where p_1 doesn't change from round to round, but the roll is independent in different rounds. These also include pure strategies ($p_1 = 0$ or $p_1 = 1$). To specify a mixed strategy, we just need a number $p_1 \in [0, 1]$.

In general, we'll use p_1 for player A's strategy, and q_1 for B's strategy.

Average expected payoff

Definition

The **average expected payoff** to a player is the average of payoffs weighted by the probability of reaching each payoff. If A plays p_1 and B plays q_1, the average expected payoff to A is called

$$v_A(p_1, q_1)$$

Example

Suppose A plays $p_1 = 0.4$ and B plays $q_1 = 0.3$.

		0.3	0.7
		1	2
0.4	1	0.12	0.28
0.6	2	0.18	0.42

$$v_A(0.4, 0.3) = (0.12)(1) + (0.28)(-1) + (0.18)(-1) + (0.42)(1) = 0.08$$

Second chance to play the game

Activity time! Now that you've had a bit of time to digest and think about the game, play again. Return to your neighbor and play 10 more rounds, making sure you still record the outcomes on paper.

Nash equilibrium

Definition

A **Nash equilibrium** of a game is a set of (mixed) strategies for each player which is "regret-free" for all players. That is, neither player can increase their value function by unilaterally changing strategy. Formally, (p, q) is an equilibrium if

$$\begin{aligned} v_A(p, q) &\geq v_A(p', q) && \text{for any alternative strategy } p' \text{ for A} \\ v_B(p, q) &\geq v_B(p, q') && \text{for any alternative strategy } q' \text{ for B} \end{aligned}$$

i.e. both players are maximizing their value function, holding strategies of other players fixed.

Equilibrium for the matching game

Proposition

The only Nash equilibrium in the matching game is $(p_1, q_1) = \left(\frac{1}{2}, \frac{1}{2}\right)$. *The average expected payoff to each player at this equilibrium is zero.*

Proof.

If either player plays 1 and 2 equally at random, then the average expected payoff to both is zero.

If either player plays 1 more than half the time, they get exploited. For example, if A plays 1 more than 2, then B exploits this by playing only 2. If A plays 1 less than 2, B exploits this by playing only 1.

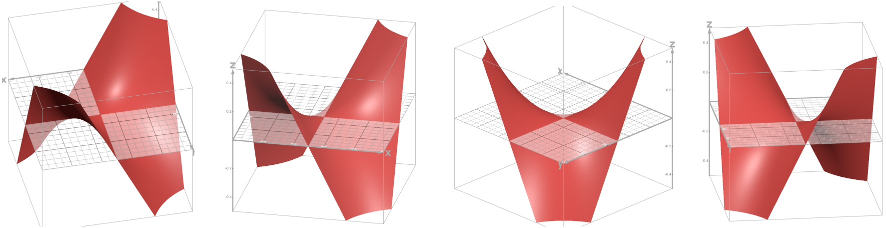
Similar reasoning if B plays 1 more or less than half the time. □

Geometric view of Nash equilibrium

Geometrically, a Nash equilibrium is a saddle point of a surface. We can graph the value function

$$v_A(p, q) = pq - p(1 - q) - q(1 - p) + (1 - p)(1 - q)$$

on the square $[0, 1] \times [0, 1]$. In equilibrium, the height is maximized in the p-direction and minimized in the q-direction, i.e. a saddle point.



The big match

We will make one small change to the game: If A ever plays 2, then both players must continue to play whatever move they played in that round, in all future rounds.

	1	2
1	1	-1
2	-1^*	1^*

Because the game could go on forever, we use average payoff per round, not the accumulated sum of round payoffs.

Round	1	···	Last 1	First 2	2 Forever	···	
A	1	···	1	**2**	*2*	*2*	···
B	?	···	?	**X**	*X*	*X*	···

Round	1	2	3	···		
A	1	1	1	···	1	···
B	?	?	?	···	?	···

Example game of the big match

In the first two rounds, A plays 1. In round 3, A plays 2, making round 3 the "big round." Since B played 2 in that critical round, they'll have to play 2 forever after. A will also play 2 in future rounds.

Round	1	2	**3**	4	5	$\cdots$
A	1	1	**2**	*2*	*2*	$\cdots$
B	2	1	**2**	*2*	*2*	$\cdots$

In a 10 round game, the payoff to A would be $+0.9$. In an infinite game, their payoff would be $+1$. If in round 3, B had played 1 instead, then the game would have gone

Round	1	2	**3**	4	5	$\cdots$
A	1	1	**2**	*2*	*2*	$\cdots$
B	2	1	**1**	*1*	*1*	$\cdots$

With 10 rounds, the payoff is -0.9 for A, or with infinite rounds -1.

Play the big match

Activity time! Play a game of big match with your partner, stopping after 20 rounds. Then switch roles and play another game.

General thoughts on the big match

Thoughts of player A

- I feel so much pressure about when to choose the critical round
- If I play 2 and B plays 1, I lose everything
- If I play 1 too many times, B will just play 1 and win many rounds

Thoughts of player B

- I feel powerless because A gets to choose when the critical round happens
- If I play 1 often to try and win the early rounds, I'm at risk of losing everything when they play 2

Questions to explore:

- Is this game really any different? Can't we still just play 1 and 2 equally often randomly and get about the same outcome?
- How should B approach playing the big match?
- How should A approach playing the big match?

Analysis of big match from B's perspective

Proposition (Blackwell and Ferguson, 1968)

If B plays the strategy $q_1 = \frac{1}{2}$, the average expected payoff for both players is zero, regardless of A's strategy.

Proof.

If A ever plays 2 (a "big round"), then in that round there is equal probability of 1 and 2 from B, so the payoff will be $+1$ half the time and -1 half the time.

If A only plays 1, then on average half the rounds will be a match and half the rounds a mismatch, resulting in an average payoff of zero (by the strong law of large numbers). □

Takeaway: *B* can play the game as if it were just the basic matching game and get the same outcome in an un-exploitable way.

Analysis of big match from A's perspective, part 1

Proposition

If A plays a strategy in which they play the same mixed strategy in every round, then there is a strategy for B which results in payoffs of $(-1, 1)$.

Proof.

If A plays 1 forever, B plays 2 forever. B wins all rounds.

If A plays some a mixed strategy $0 \leq p_1 < 1$, then B plays 1 forever. Eventually, A will play 2, and B will win all future rounds. □

Takeaway: If A plays big match as if it were just the basic matching game, they are very exploitable. In particular, playing 1 half the time in all rounds is a losing strategy if B responds appropriately.

History-based strategies

Definition

In a game repeated over multiple rounds, a **stationary strategy** is the same in each round (but can be mixed). In contrast, a **non-stationary** or **history-based strategy** adapts to the history of previous moves.

Example

"Play 1 half the time" is a stationary strategy. Examples of history-based strategies include:

- Play the opposite of my opponent's move in the previous round.
- Play 1 40% of the time, then 30% of the time, then 60% of the time, then repeat this.
- If my opponent has played 1 x times in all n rounds played, I will play 1 with probability x/n.

Analysis of big match from A's perspective, part 2

Proposition (Lester Dubins)

There is no strategy for A which obtains average expected payoff zero or more against every strategy of B.

Proof. (by contradiction)

Suppose not, i.e. A has a strategy S that averages at least zero against any possible strategy for B. S cannot be playing 1 always, since B exploits this by playing 2 always, giving A the payoff -1.

(continued on next slide)

Let m be the first round in which A has probability $\varepsilon > 0$ of playing 2.

Round	1	$\cdots$	$m-1$	m	$\cdots$
A : Strategy S	1	$\cdots$	1	$p_2 = \varepsilon > 0$	$\cdots$

But B has a strategy S' which defeats strategy S: 2 for rounds 1 through $m-1$, 1 in round m, then $q_1 = \frac{1}{2}$ in forever after.

Round	1	$\cdots$	$m-1$	m	$m+1$	$\cdots$
A : Strategy S	1	$\cdots$	1	$p_2 = \varepsilon > 0$	?	$\cdots$
B : Strategy S'	2	$\cdots$	2	1	$q_1 = \frac{1}{2}$	$\cdots$

If A ends up playing 2 in round m, then they lose the "big round" and get payoff -1. If they play 1 in round m, they get average payoff zero from the long tail. Overall, A's average expected payoff from both players implementing these strategies is $-\varepsilon < 0$.

Takeaway: No strategy for A is completely exploit-proof.

A strategy for player A

Definition

Fix a non-negative integer N. As player A, after n rounds (of playing 1), count the number of times B has played 1 and 2. Calculate

$$E_2 = \text{excess 2's} = (\#2 \text{ by B}) - (\#1 \text{ by B})$$

In round $n+1$, play 2 with probability

$$p_2(E_2) = \frac{1}{(N - E_2 + 1)^2} = \frac{1}{(N + \#1 - \#2 + +1)^2}$$

This strategy is called $S(N)$.

Intuition: Make N large, so playing 2 is rare. If B plays 1 a lot, play 1 even more. If B plays 2 a lot, play 2 slightly more often.

$N = 9$

Example

In round 1, there are no previous rounds so $E_2 = 0$. So in round 1,

$$p_2(0) = \frac{1}{(9-0+1)^2} = \frac{1}{10^2} = \frac{1}{100}$$

After 49 rounds, B has played 29 1's and 20 2's. Then $E_2 = -9$, so in round 50 we play 2 even less than 1%.

$$p_2(9) = \frac{1}{(9-(-9)+1)^2} = \frac{1}{19^2} = \frac{1}{361}$$

If B had played 20 1's and 29 2's, then $E_2 = 9$ so in round 50 we play 2 guaranteed.

$$p_2(-9) = \frac{1}{(9-9+1)^2} = \frac{1}{1^2} = 1$$

How good is $S(N)$ for A?

Theorem (Blackwell and Ferguson, 1968)

If A plays strategy $S(N)$, then their average expected payoff (regardless of what B does) is at least $\dfrac{-1}{2N+1}$. So for $\varepsilon > 0$, A can choose N sufficiently large that $v_A \geq -\varepsilon$.

Proof.

- If B plays a sequence of moves in which E_2 eventually reaches N (i.e B plays a lot more 2's than 1's), then the game is guaranteed to end in a finite number of rounds. (Then do induction.)
- If B plays a sequence of moves in which E_2 never reaches N (i.e. B never plays more than N more 2's than 1's), then ratio of 2's to 1's is never too far from $\frac{1}{2}$, so the average expected payoff can't get too far from zero.

$\square$

Final analysis of big match

Corollary

Neither player can guarantee better than an average expected payoff of zero in the big match.

Proof.

From earlier, B can flip a coin each round and get an average of zero.

If B could do better, i.e. guarantee themselves a strictly positive number $\varepsilon > 0$, this would mean that A is guaranteed to get less than $-\varepsilon$, but we know that A can use $S(N)$ to get at least $-\varepsilon$. So it can't be possible for B to do this. □

Final summary: If B plays 1 half the time, they can't be exploited at all. If A chooses $\varepsilon > 0$ and then N large enough then they can only lose at most $-\varepsilon$ on average. This is not a Nash equilibrium, but neither player can improve by more than ε by changing strategies.

Repeated games with absorbing states

Suppose two players repeatedly play a zero-sum matrix game, with some absorbing entries. Let v_n be the average expected payoff (to player A) from optimal play over n rounds.

Theorem (Kolhberg, 1974)

- $\lim\limits_{n\to\infty} v_n$ *converges*
- *For every* $\varepsilon > 0$, *each player has an* ε*-optimal strategy which results in an average expected payoff within* ε *of some number* v_∞
- *If information is symmetric and players have perfect recall,* $\lim\limits_{n\to\infty} v_n = v_\infty$

Example

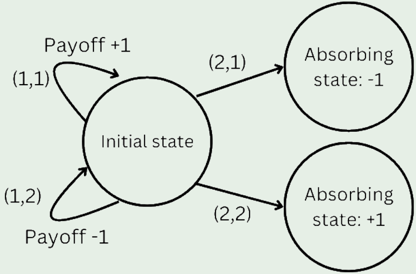
In the big match, $v_n = 0$ regardless of n, so $\lim\limits_{n\to\infty} v_n = v_\infty = 0$.

Stochastic games

Definition

A **stochastic game** has states, players, and actions. In each round, players choose an action, players get payoffs, then the next state is probabilistically determined by the actions and the previous state.

Example (The big match)



Uniform value

Definition

In a stochastic game, if for any $\varepsilon > 0$, a player has a strategy guaranteeing payoff at least $v - \varepsilon$ (and they can't guantee more), then v is the **uniform value** of the game to that player. Such strategies are called ε-**optimal strategies**.

Theorem (Mertens & Neyman, 1979)

In every stochastic game with finitely many states and players, every player has a uniform value, i.e. all players have ε-optimal strategies.

Example

The uniform value of the big match is zero to both players. B has a strategy which is 0-optimal, while A only has ε-optimal strategies.

Research directions in stochastic games

- What if the state space is infinite/continuous?
- What if players have incomplete or asymmetrical information?
- What if players don't have perfect recall of the game's history?
- Even if theoretical approaches are known for finding good strategies in stochastic games, are there computationally efficient ways to find and implement good strategies?