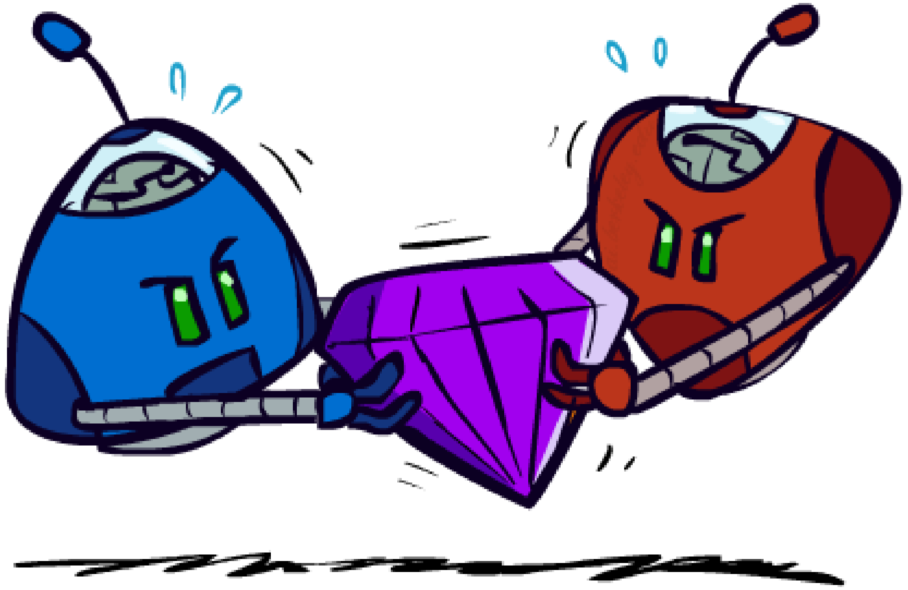
Announcements

Assignments (everything left for the semester):

- P5 due Thu 4/30
- HW12 (written) due Tue 4/28

AI: Representation and Problem Solving

Game Theory: Equilibrium



Instructors: Pat Virtue & Stephanie Rosenthal

Slide credits: CMU AI

Image credit: ai.berkeley.edu

Learning Objectives

Formulate a problem as a game

Describe and compare the basic concepts in game theory

- Normal-form game, extensive-form game
- Zero-sum game, general-sum game
- Pure strategy, mixed strategy, support, best response, dominance
- Dominant strategy equilibrium, Nash equilibrium, (Minimax strategy), (maximin strategy), Stackelberg equilibrium

Describe iterative removal algorithm

(Describe minimax theorem)

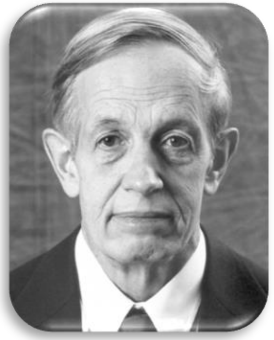
Compute equilibria for bimatrix games

- Pure strategy Nash equilibrium
- Mixed strategy Nash equilibrium (including using LP for zero-sum games)
- Stackelberg equilibrium (only pure strategy equilibrium is required)

From Games to Game Theory

The study of mathematical models of conflict and cooperation between intelligent decision makers

Used in economics, political science, etc

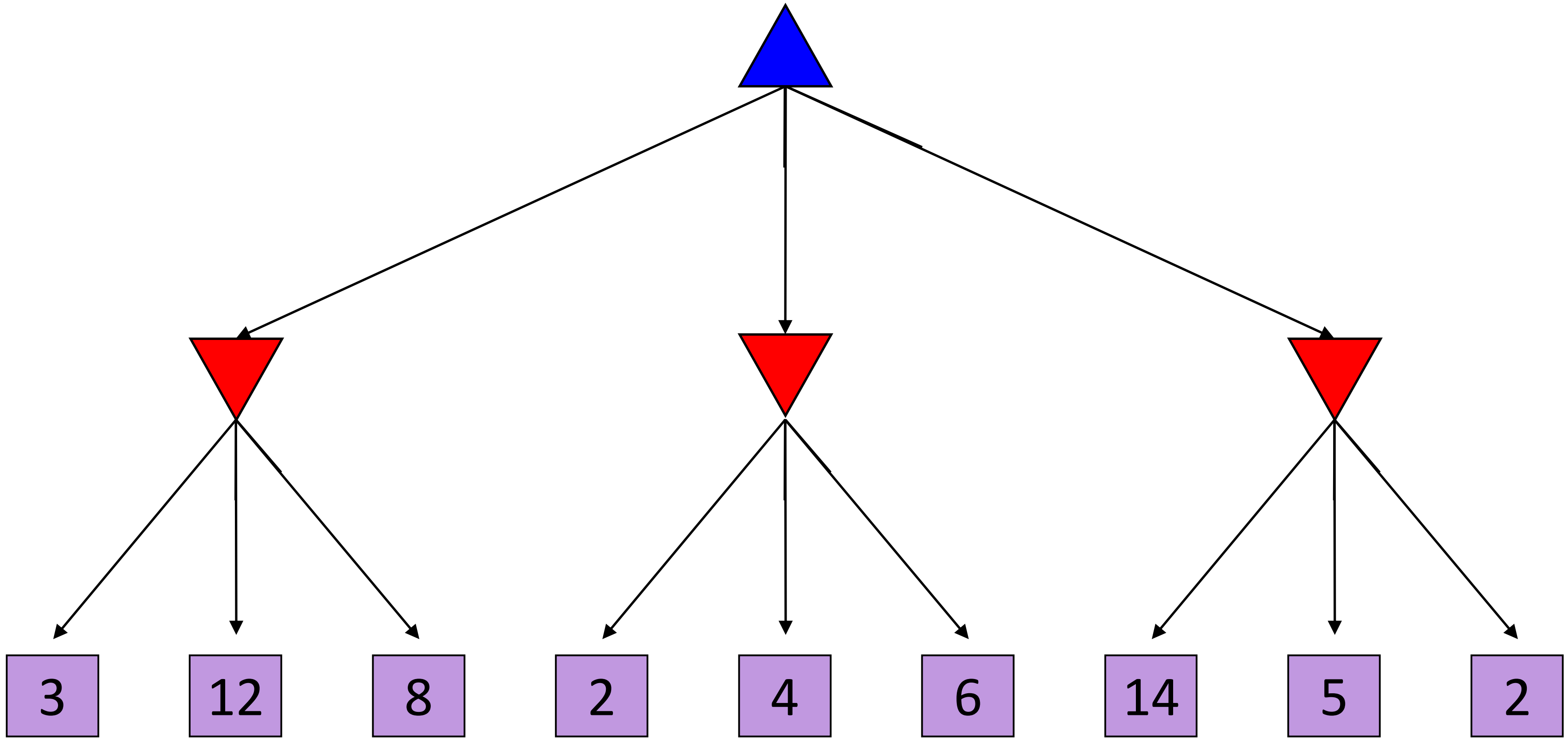


John Nash

Winner of Nobel Memorial Prize in Economic Sciences

Recall: Adversarial Search

Zero-sum, perfect information, two player games with turn-taking moves



Classical Games and Payoff Matrices

Rock-Paper-Scissors (RPS)

- Rock beats Scissors
- Scissors beats Paper
- Paper beats Rock

Player 2

Player 1

	Rock	Paper	Scissors
Rock	0, 0	-1, 1	1, -1
Paper	1, -1	0, 0	-1, 1
Scissors	-1, 1	1, -1	0, 0

Rock, Paper, Scissors, Lizard, Spock

CBS, Big Bang Theory

https://www.youtube.com/watch?v=iSHPVCBsnLw

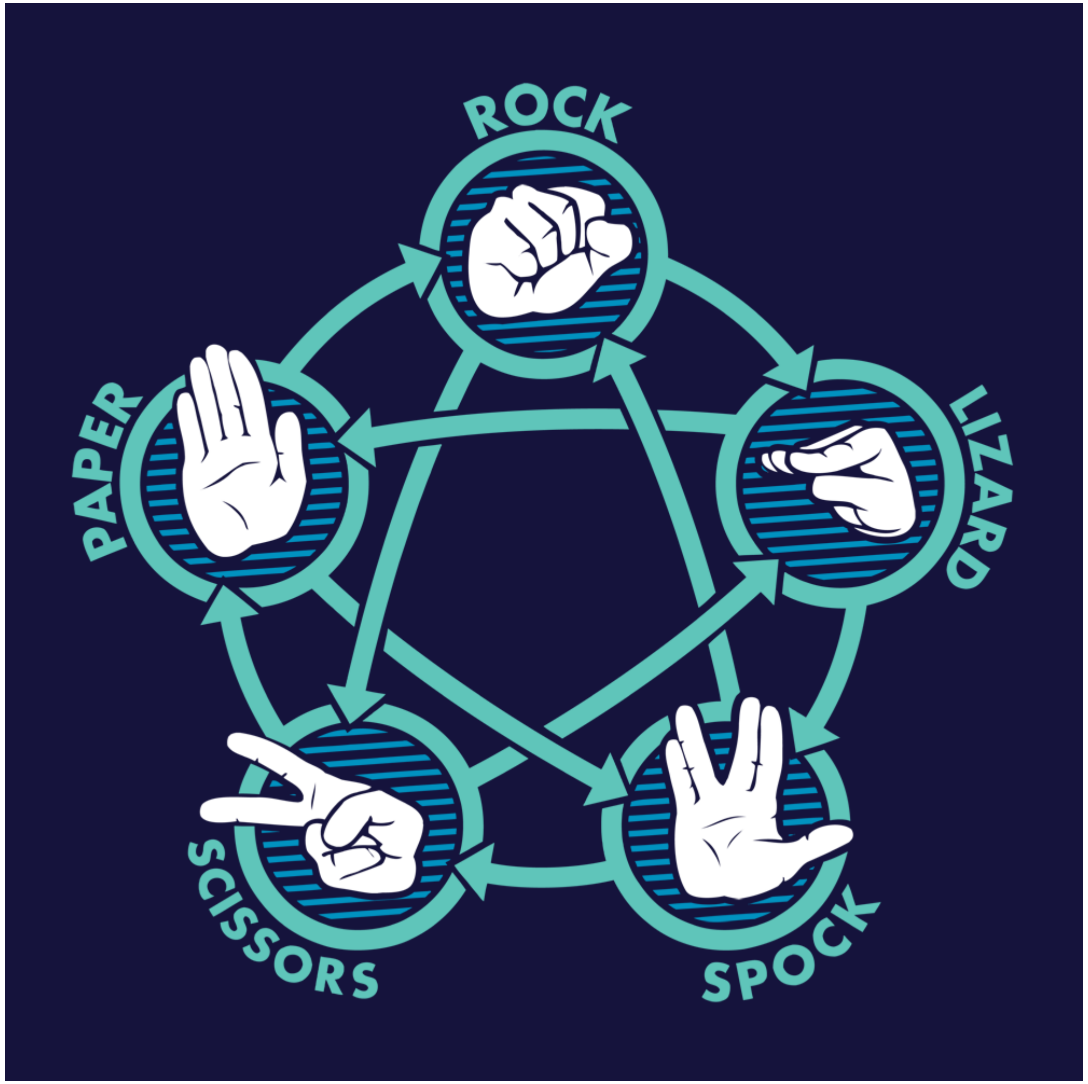


Image credit: https://www.snorgtees.com/rock-paper-scissors-lizard-spock

Classical Games and Payoff Matrices

Prisoner’s Dilemma (PD)

- If both Cooperate: 1 year in jail each
- If one Defect, one Cooperate: 0 year for (D), 3 years for (C)
- If both Defect: 2 years in jail each
- Let’s play!

Player 2

Player 1

	Cooperate	Defect
Cooperate	-1,-1	-3,0
Defect	0,-3	-2,-2

Variation: Split or Steal

https://youtu.be/p3Uos2fzIJ0

Classical Games and Payoff Matrices

Football vs Concert (FvsC)

- ~~Historically known as Battle of Sexes~~
- If football together: Alex ☺☺, Berry ☺
- If concert together: Alex ☺, Berry ☺☺
- If not together: Alex ☹, Berry ☹

utilities

Fill in the payoff matrix

Berry

Alex		Football	Concert
	Football	2, 1	0, 0
	Concert	0, 0	1, 2

Normal-Form Games

A game in normal form consists of the following elements

- Set of players
- Set of actions for each player
- Payoffs / Utility functions
  - Determines the utility for each player given the actions chosen by all players (referred to as action profile)
- Bimatrix game is special case: two players, finite action sets

Players move simultaneously and the game ends immediately afterwards

Strategy: Rock, Paper, Scissors

Design an AI to play Rock-Paper-Scissors for T rounds

If T=1, roundID=1, what action does your AI choose?

```
function playRPS(roundID,T)
Input: $T$, roundID ∈ $\{1..T\}$
Output: action a ∈ {Rock, Paper, Scissors}
```

mixed	1/3	1/3	1/3
pure	1	0	0

```
Return a
```

Strategy

Pure strategy: choose an action deterministically

Mixed strategy: choose actions according to a probability distribution

- Notation: $s = (\underline{0.3}, \underline{0.7}, 0)$
- Support: set of actions chosen with non-zero probability

Notation Alert! We use s to represent strategy here (not states)

Does your AI play a deterministic strategy or a mixed strategy?

What is the support size of your AI's strategy?

3

Player 2

Player 1

		Rock	Paper	Scissors
$1/3$	Rock	0, 0	-1, 1	1, -1
$1/3$	Paper	1, -1	0, 0	-1, 1
$1/3$	Scissors	-1, 1	1, -1	0, 0

Zero-sum vs General-sum

Zero-sum Game

- No matter what actions are chosen by the players, the utilities for all the players sum up to zero or a constant

General-sum Game

- The sum of utilities of all the players is not a constant

Which ones are general-sum games?

Player 2

Player 1

	Cooperate	Defect
Cooperate	-1,-1	-3,0
Defect	0,-3	-2,-2

Player 2

Player 1

	Rock	Paper	Scissors
Rock	0,0	-1,1	1,-1
Paper	1,-1	0,0	-1,1
Scissor	-1,1	1,-1	0,0

Berry

Alex

	Football	Concert
Football	2,1	0,0
Concert	0,0	1,2

Expected Utility

Notation Alert!

Use a, s, u to represent action, strategy, utility of a player

Use $\mathbf{a}, \mathbf{s}, \mathbf{u}$ to represent action, strategy, utility profile (a set of players)

Given the strategies of all players,

Expected Utility for player i u_i=

$$\sum_{\mathbf{a}} \text{Prob(action profile } \mathbf{a}) \times \text{Utility for player } i \text{ in } \mathbf{a}$$

Can skip action profiles with probability 0 or utility 0

If Alex's strategy $s_A = \left(\frac{1}{2}, \frac{1}{2}\right)$, Berry's strategy $s_B = (1,0)$

What is the probability of action profile $\mathbf{a} = (\underset{1}{\text{Concert}}, \underset{2}{\text{Football}})$?

$$1/2 \cdot 1 = 1/2$$

What is Alex's utility in this action profile?

$$u_A(C, F) = 0$$

Berry

		1	0
		Football	Concert
Alex	1/2 Football	2,1	0,0
	1/2 Concert	$\boxed{0,0}$	1,2

Piazza Poll 1

Expected Utility for player i u_i=

$$\sum_{\mathbf{a}} \text{Prob(action profile } \mathbf{a}) \times \text{Utility for player } i \text{ in } \mathbf{a}$$

In Rock-Paper-Scissors, if $s_1 = \left(\frac{1}{3}, \frac{2}{3}, 0\right)$, $s_2 = \left(0, \frac{1}{2}, \frac{1}{2}\right)$, how many non-zero terms need to be added up when computing the expected utility for player 1?

A. 9

B. 6

C. 4

🚫 ~~D. 7~~

[✓ E. 3]

Player 2

Player 1

		0	1/2	1/2
		~~Rock~~	Paper	Scissors
1/3	Rock	~~0, 0~~	-1, 1	1, -1
2/3	Paper	~~1, -1~~	~~0, 0~~	-1, 1
0	~~Scissors~~	~~-1, 1~~	~~1, -1~~	~~0, 0~~

Piazza Poll 2

Expected Utility for player i u_i=

$$\sum_{\mathbf{a}} \text{Prob(action profile } \mathbf{a}) \times \text{Utility for player } i \text{ in } \mathbf{a}$$

In Rock-Paper-Scissors, if $s_1 = \left(\frac{1}{3}, \frac{2}{3}, 0\right)$, $s_2 = \left(0, \frac{1}{2}, \frac{1}{2}\right)$, what is the utility of player 1?

A. -1

[✓ B. -1/3]

C. 0

D. 2/3

🚫 E. 1

			Player 2	
		0	1/2	1/2
		~~Rock~~	Paper	Scissors
Player 1 1/3	Rock	~~0, 0~~	-1, 1	1, -1
2/3	Paper	~~1, -1~~	~~0, 0~~	-1, 1
0	~~Scissors~~	~~-1, 1~~	~~1, -1~~	~~0, 0~~

$u_1(R,P) = \underbrace{\frac{1}{3}\frac{1}{2}} \cdot -1 = -\frac{1}{6}$

$u_1(R,S) = \frac{1}{3}\frac{1}{2}\ 1 = \frac{1}{6}$

$u_1(P,S) = \frac{2}{3}\frac{1}{2} -1 = -\frac{2}{6}$

$-\frac{1}{3}$

Best Response

Best Response (BR): Given the strategies or actions of all players but player i (denoted as $\mathbf{s}_{-i}$ or $\mathbf{a}_{-i}$), Player i's best response to $\mathbf{s}_{-i}$ or $\mathbf{a}_{-i}$ is the set of actions or strategies of player i that can lead to the highest expected utility for player i

$$u_1(s_1, s_{-1}) \geq u_1(s_1', s_{-1}) \; \forall s_1'$$

In RPS, what is Player 1's best response to Rock (i.e., assuming Player 2 plays Rock)?

$P = BR(R)$

$= BR(\{1/3, 1/3, 1/3\})$

In Prisoner's Dilemma, what is Player 1's best response to Cooperate? What is Player 1's best response to Defect?

Defect

Defect

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	-1,-1	-3,0
	Defect	0,-3	-2,-2

Best Response

Best Response (BR): Given the strategies or actions of all players but player i (denoted as $\mathbf{s}_{-i}$ or $\mathbf{a}_{-i}$), Player i's best response to $\mathbf{s}_{-i}$ or $\mathbf{a}_{-i}$ is the set of actions or strategies of player i that can lead to the highest expected utility for player i

In RPS, what is Player 1's best response to Rock (i.e., assuming Player 2 plays Rock)?

Paper

In Prisoner's Dilemma, what is Player 1's best response to Cooperate? What is Player 1's best response to Defect?

Defect, Defect

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	-1,-1	-3,0
	Defect	0,-3	-2,-2

Best Response

Best Response (BR): Given the strategies or actions of all players but player i (denoted as $\mathbf{s}_{-i}$ or $\mathbf{a}_{-i}$), Player i's best response to $\mathbf{s}_{-i}$ or $\mathbf{a}_{-i}$ is the set of actions or strategies of player i that can lead to the highest expected utility for player i

What is Alex's best response to Berry's mixed strategy $s_B = \left(\frac{1}{2}, \frac{1}{2}\right)$?

$$u_A(F, s_B) = \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 0 = 1$$

$$u_A(C, s_B) = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 1 = \frac{1}{2}$$

		Berry	
		1/2	1/2
		Football	Concert
Alex	Football	$\underline{2}$,1	$\underline{0}$,0
	Concert	$\underline{0}$,0	$\underline{1}$,2

Piazza Poll 3

In Rock-Paper-Scissors, if $s_1 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, which actions or strategies are player 2's best responses to s_1?

[✓ A.] Rock

[✓ B.] Paper

[✓ C.] Scissors

🚫 D. Lizard

[✓ E.] $s_2 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$

[✓ F.] $s_2 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$

$u_2(s_1, R) = \frac{1}{3} \cdot 0 + \frac{1}{3}(-1) + \frac{1}{3} \cdot 1 = 0$

$u_2(s_1, P) = 0$

$u_2(s_1, S) = 0$

			Player 2		
			1/3	1/3	1/3
			Rock	Paper	Scissors
1/3	Player 1	Rock	0, 0	-1, 1	1, -1
1/3		Paper	1, -1	0, 0	-1, 1
1/3		Scissors	-1, 1	1, -1	0, 0

Best Response

Theorem 1 (Nash 1951): A mixed strategy is BR iff all actions in the support are BR

Player 2

Player 1

	Rock	Paper	Scissors
Rock	0, 0	-1, 1	1, -1
Paper	1, -1	0, 0	-1, 1
Scissors	-1, 1	1, -1	0, 0

Dominance

s_i and $s_i{}'$ are two strategies for player i

s_i strictly dominates $s_i{}'$ if s_i is always better than $s_i{}'$, no matter what strategies are chosen by other players

s_i strictly dominates s_i' if

$$u_i\ (\underline{s_i}, \mathbf{s}_{-i}) > u_i(\underline{s_i'}, \mathbf{s}_{-i}), \underline{\forall \mathbf{s}_{-i}}$$

always better

s_i very weakly dominates s_i' if

$$u_i\ (s_i, \mathbf{s}_{-i}) \geq u_i(s_i', \mathbf{s}_{-i}), \forall \mathbf{s}_{-i}$$

never worse

s_i weakly dominates s_i' if

$$u_i\ (s_i, \mathbf{s}_{-i}) \geq u_i(s_i', \mathbf{s}_{-i}), \forall \mathbf{s}_{-i}$$

and $\exists \mathbf{s}_{-i}, u_i\ (s_i, \mathbf{s}_{-i}) > u_i(s_i', \mathbf{s}_{-i})$

never worse and sometimes better

Dominance

Can you find any dominance relationships between the pure strategies in these games?

D > C

Player 2

Player 1

	Rock	Paper	Scissors
Rock	0,0	-1,1	1,-1
Paper	1,-1	0,0	-1,1
Scissor	-1,1	1,-1	0,0

Player 2

Player 1

	Cooperate	Defect
Cooperate	-1,-1	-3,0
Defect	0,-3	-2,-2

Berry

Alex

	Football	Concert
Football	2,1	0,0
Concert	0,0	1,2

Dominance

If s_i strictly dominates $s_i', \forall s_i' \in S_i \backslash \{s_i\}$,

is s_i a best response to $\mathbf{s}_{-i}, \forall \mathbf{s}_{-i}$?

Yes. Remember:

- s_i strictly dominates s_i' if

$$u_i\,(s_i, \mathbf{s}_{-i}) > u_i(s_i', \mathbf{s}_{-i}), \forall \mathbf{s}_{-i}$$

Rewriting the statement at the top:

$$u_i\,(s_i, \mathbf{s}_{-i}) > u_i(s_i', \mathbf{s}_{-i}), \forall \mathbf{s}_{-i} \quad \forall s_i' \in S_i \backslash \{s_i\}$$

So… for any $\mathbf{s}_{-i}$

$$u_i\,(s_i, \mathbf{s}_{-i}) > u_i(s_i', \mathbf{s}_{-i}), \forall s_i' \in S_i \backslash \{s_i\}$$

This is the definition of best response ☺

That is, s_i leads to the highest utility compared to all other responses, $s_i{}'$

Solution Concepts in Games

How should one player play and what should we expect all the players to play?

- 1 Dominant strategy and dominant strategy equilibrium
- 2 Nash Equilibrium
- (Minimax strategy)
- (Maximin strategy)
- 3 Stackelberg Equilibrium

Dominant Strategy

A strategy could be (always better / never worse / never worse and sometimes better) than any other strategy

s_i is a (strictly/very weakly/weakly) dominant strategy if it (strictly/very weakly/weakly) dominates $s_i', \forall s_i' \in S_i \setminus \{s_i\}$

Focus on single player’s strategy

Doesn’t always exist

Is there a strictly dominant strategy for player 1 in PD?

Player 2

		Cooperate	Defect
Player 1	Cooperate	-1,-1	-3,0
	Defect	0,-3	-2,-2

Dominant Strategy Equilibrium

Sometimes called dominant strategy solution

Every player plays a dominant strategy

Focus on strategy profile for all players

Doesn't always exist

What is the dominant strategy equilibrium for PD?

Player 2

Player 1

	Cooperate	Defect
Cooperate	-1,-1	-3,0
Defect	0,-3	-2,-2

Solution Concepts in Games

How should one player play and what should we expect all the players to play?

- Dominant strategy and dominant strategy equilibrium
- Nash Equilibrium
- (Minimax strategy)
- (Maximin strategy)
- Stackelberg Equilibrium

Nash Equilibrium

Nash Equilibrium (NE)

- Every player’s strategy is a best response to others’ strategy profile
- In other words, one cannot gain by unilateral deviation
- Pure Strategy Nash Equilibrium (PSNE)
  - $\underline{a_i} \in BR(\mathbf{a}_{-i}), \forall i$
- Mixed Strategy Nash Equilibrium
  - At least one player use a randomized strategy
  - $\underline{s_i} \in BR(\mathbf{s}_{-i}), \forall i$

Nash Equilibrium

What are the PSNEs in these games?

D, D

F, C C, F

What is the mixed strategy NE in RPS?

1/3, 1/3, 1/3

Player 2

Player 1

	Rock	Paper	Scissors
Rock	0,0	-1,1	1,-1
Paper	1,-1	0,0	-1,1
Scissor	-1,1	1,-1	0,0

Player 2

Player 1

	Cooperate	Defect
Cooperate	-1,-1	-3,0
Defect	0,-3	-2,-2

Berry

Alex

	Football	Concert
Football	2,1	0,0
Concert	0,0	1,2

Nash Equilibrium

Theorem 2 (Nash 1951): NE always exists in finite games

- Finite number of players, finite number of actions
- NE: can be pure or mixed
- Proof: Through Brouwer's fixed point theorem

Find PSNE

Find pure strategy Nash Equilibrium (PSNE)

- Enumerate all action profile
- For each action profile, check if it is NE
  - For each player, check other available actions to see if he should deviate
- Other approaches?

Player 2

		L	C	R
Player 1	U	10, 3	1, 5	5, 4
	M	3, 1	2, 4	5, 2
	D	0, 10	1, 8	7, 0

Find PSNE

A strictly dominated strategy is one that is always worse than some other strategy

Strictly dominated strategies cannot be part of an NE **Why?**

Which are the strictly dominated strategies for player 1?

∅

How about player 2?

R

Player 2

Player 1	L	C	R
U	10, 3	1, 5	5, 4
M	3, 1	2, 4	5, 2
D	0, 10	1, 8	7, 0

Find PSNE through Iterative Removal

Remove strictly dominated actions (pure strategies) and then find PSNE in the remaining game

Can have new strictly dominated actions in the remaining game!

Repeat the process until no actions can be removed

This is the Iterative Removal algorithm (also known as Iterative Elimination of Strictly Dominated Strategies)

Find PSNE in this game using iterative removal

Player 2

Player 1

	~~L~~	C	~~R~~
U	~~10, 3~~	~~1, 5~~	~~5, 4~~
M	~~3, 1~~	$\boxed{2, 4}$	~~5, 2~~
~~D~~	~~0, 10~~	~~1, 8~~	~~7, 0~~

Find Mixed Strategy Nash Equilibrium

Can we still apply iterative removal?

- Yes! The removed strategies cannot be part of any NE
- You can always apply iterative removal first

Find Mixed Strategy Nash Equilibrium

How to find mixed strategy NE (after iterative removal)?

q Berry $1-q$

			Football	Concert
p	Alex	Football	2,1	0,0
$1-p$		Concert	0,0	1,2

If $s_A = (p, 1-p)$ and $s_B = (q, 1-q)$ with $0 < p, q < 1$ is a NE, what are the necessary conditions for p and q?

Find Mixed Strategy Nash Equilibrium

How to find mixed strategy NE (after iterative removal)?

$\frac{1}{3}$ **Berry** $\frac{2}{3}$

Alex		Football	Concert
ε $\frac{2}{3}$	Football	$\underline{2}$,1	0,0
$1-\varepsilon$ $\frac{1}{3}$	Concert	0,0	$\underline{1}$,2

Is $s_A = \left(\frac{2}{3}, \frac{1}{3}\right)$ and $s_B = \left(\frac{1}{3}, \frac{2}{3}\right)$ an NE?

$$\underline{u_A(s_A, s_B)} \geq u_A((\varepsilon, 1-\varepsilon), s_B)$$

$$\underline{\tfrac{2}{3}}\cdot\tfrac{1}{3}\cdot 2 + \underline{\tfrac{1}{3}}\cdot\tfrac{2}{3}\cdot 1 = \varepsilon\cdot\tfrac{1}{3}\cdot 2 + (1-\varepsilon)\cdot\tfrac{2}{3}\cdot 1$$

$$\tfrac{2}{3} \; \underline{\underline{=}} \; \cancel{\varepsilon\tfrac{2}{3}} + \tfrac{2}{3} - \cancel{\varepsilon\tfrac{2}{3}}$$

$$\to \varepsilon \underbrace{u_A(F, s_B)}_{v} + (1-\varepsilon)\underbrace{u_A(C, s_B)}_{v}$$

$$= \varepsilon v + (1-\varepsilon) v$$

$$= v$$

<u>Want</u>

$$u_A(F, s_B) = u_A(C, s_B)$$

Find Mixed Strategy Nash Equilibrium

How to find mixed strategy NE (after iterative removal)?

Berry

		Football	Concert
Alex	Football	2,1	0,0
	Concert	0,0	1,2

If $s_A = (p, 1 - p)$ and $s_B = (q, 1 - q)$ with $0 < p, q < 1$ is a NE, what are the necessary conditions for p and q?

Find Mixed Strategy Nash Equilibrium

How to find mixed strategy NE (after iterative removal)?

Berry

		Football	Concert
Alex	Football	2,1	0,0
	Concert	0,0	1,2

If $s_A = (p, 1-p)$ and $s_B = (q, 1-q)$ with $0 < p, q < 1$ is a NE, what are the necessary conditions for p and q?

$$u_A(\underline{F}, s_B) = u_A(\underline{C}, s_B) \qquad u_B(s_A, F) = u_B(s_A, C)$$

Why? Remember Theorem 1: A mixed strategy is BR iff all actions in the support are BR.

So…if $s_A \in BR(s_B)$, then $F_A \in BR(s_B)$and $C_A \in BR(s_B)$

Find Mixed Strategy Nash Equilibrium

How to find mixed strategy NE (after iterative removal)?

q Berry $1-q$

Alex

	Football	Concert
Football	2,1	0,0
Concert	0,0	1,2

$s_A(2/3, 1/3)$

$s_B(1/3, 2/3)$

If $s_A = (p, 1-p)$ and $s_B = (q, 1-q)$ with $0 < p, q < 1$ is a NE, what are the necessary conditions for p and q?

$$\underline{u_A(F, s_B)} = \underline{u_A(C, s_B)}$$

$$q \cdot 2 + (1-q) \cdot 0 = q \cdot 0 + (1-q) \cdot 1$$

$$2q = 1-q$$

$$q = 1/3$$

$$\underline{u_B(s_A, F)} = \underline{u_B(s_A, C)}$$

$$p \cdot 1 + (1-p) \cdot 0 = p \cdot 0 + (1-p) \cdot 2$$

$$p = 2 - 2p$$

$$p = 2/3$$

Piazza Poll 4

If $s_A = (p, 1-p)$ and $s_B = (q, 1-q)$ with $0 < p, q < 1$ is a NE of the game, which equations should p and q satisfy?

[✓ A.] $2q = \underline{3}(1-q)$

B. $2p = 3(1-p)$

C. $q = 2(1-q)$

[✓ D.] $p = 2(1-p)$

🚫 E. $p = q$

		Berry	
		Football	Concert
Alex	Football	2,1	0,0
	Concert	0,0	$\underline{\underline{3}}$,2

Solution Concepts in Games

How should one player play and what should we expect all the players to play?

- Dominant strategy and dominant strategy equilibrium
- Nash Equilibrium
- Minimax strategy
- Maximin strategy
- Stackelberg Equilibrium

Power of Commitment

What's the PSNEs in this game and the players' utilities? 2,1

What action should player 2 choose if player 1 commits to playing b? What is player 1's utility?

d 3

What action should player 2 choose if player 1 commits to playing a and b uniformly randomly? What is player 1's expected utility?

d 3.5

Player 2

			c	d
1/2	Player 1	a	2,1	4,0
1/2		b	1,0	3,2

Stackelberg Equilibrium

Stackelberg Game

- Leader commits to a strategy first
- Follower responds after observing the leader’s strategy

Stackelberg Equilibrium

- Follower best responds to leader’s strategy
- Leader commits to a strategy that maximize her utility assuming follower best responds

Player 2

Player 1

	c	d
a	2,1	4,0
b	1,0	3,2

Stackelberg Equilibrium

If the leader can only commit to a pure strategy, or you know that the leader's strategy in equilibrium is a pure strategy, the equilibrium can be found by enumerating leader's pure strategy

If ties for the follower are broken by the follower such that the leader benefits, the leader can exploit this. This is the strong Stackelberg equilibrium (SSE)

In general, the leader can commit to a mixed strategy and

$u^{\underline{SSE}} \geq u^{NE}$ (first-mover advantage)!

Berry

		Football	Concert
Alex	Football	2,1	0,0
	Concert	0,0	1,2

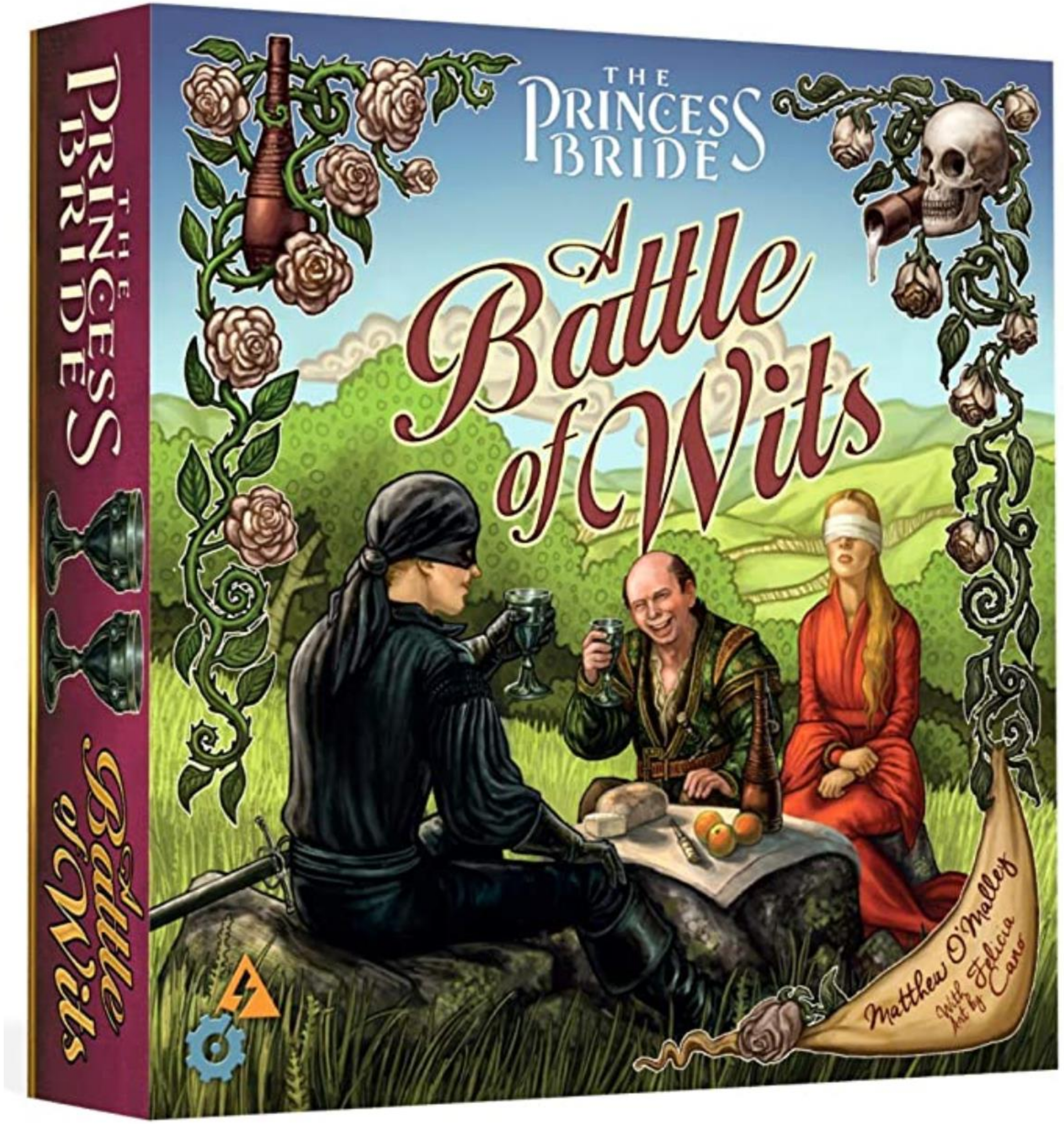
Player 2

		c	d
Player 1	a	2,1	4,0
	b	1,0	3,2

A Battle of Wits

The Princess Bride

https://www.youtube.com/watch?v=EZSx3zNZOaU



https://www.amazon.com/Princess-Bride-Battle-Wits-3rd