

# Play Rock Paper Scissors with the people around you

What are the rules?

What is the best action to take if the other player plays Rock?

What is the best overall strategy?

# Announcements

## Assignments:

- HW10 due Tuesday
- P5 due Thursday (week from today)

R14 - Last 1/3 class

## Review Sessions:

- Sunday 4/30 5-7pm in NSH 3305 - M1
- Tuesday 5/2 5-7pm in NSH 3305 - M2

## Final Exam:

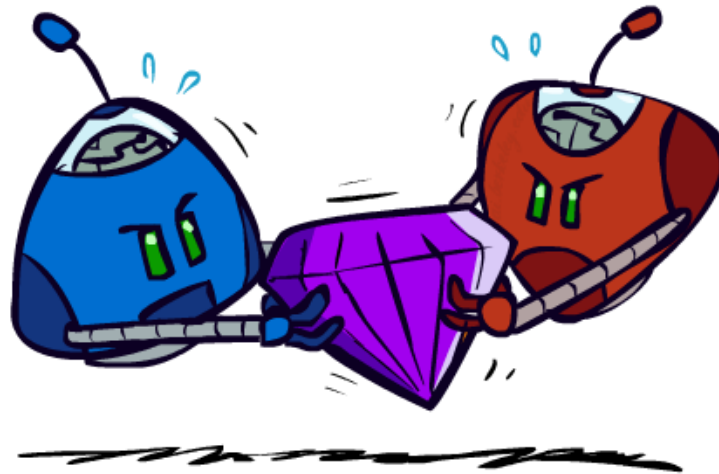
- **Thursday 5/4 5:30pm in Rashid Auditorium**
- 1/2 of exam on last 1/3 of class, all material is fair game

## Other:

- Come to office hours for conceptual questions!

# AI: Representation and Problem Solving

## Game Theory: Equilibrium



Instructors: Stephanie Rosenthal

Slide credits: CMU AI

# 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

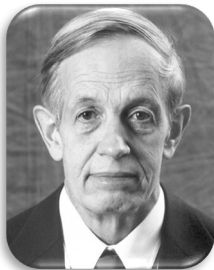
- Pure strategy Nash equilibrium
- Mixed strategy Nash equilibrium (including using LP for zero-sum games)
- Stackelberg equilibrium (only pure strategy equilibrium is required)

# From Games to Game Theory



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

Used in economics, political science, etc

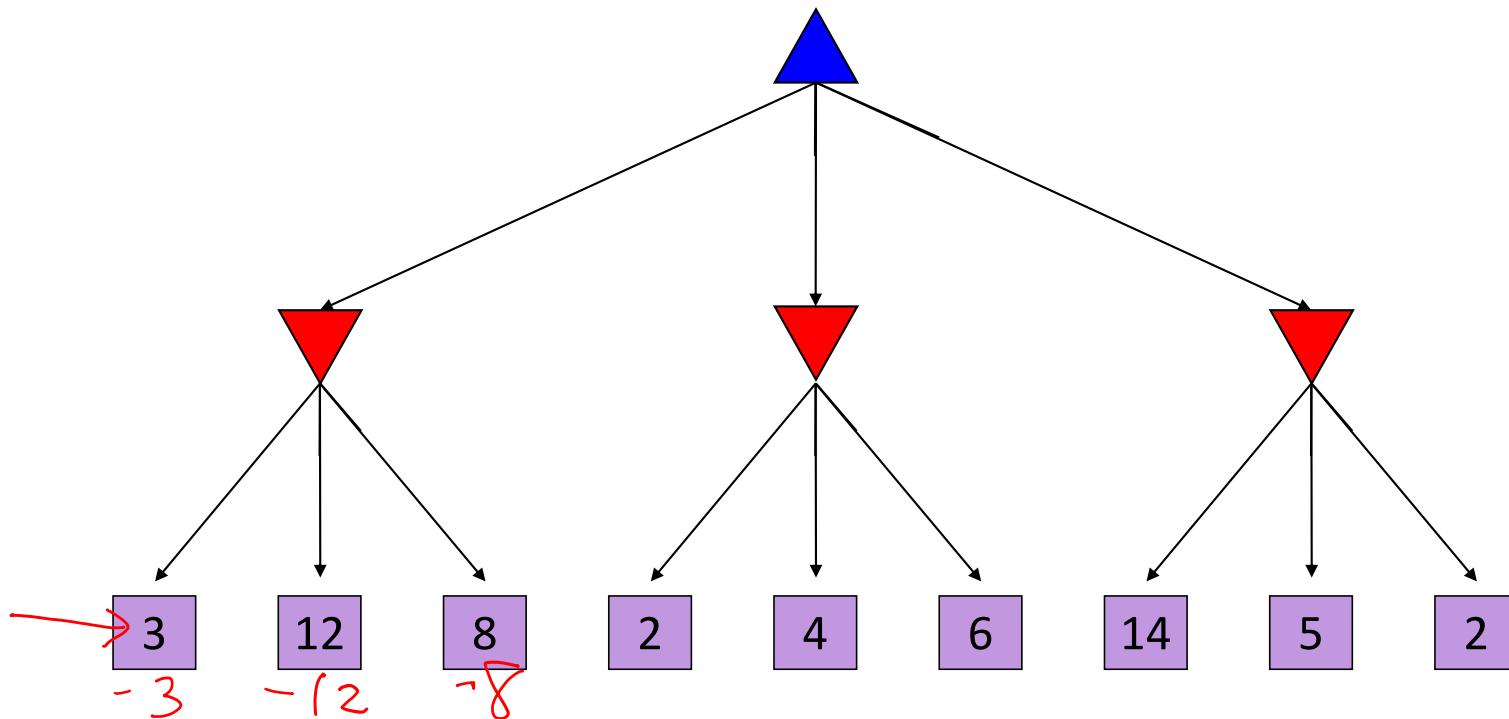


John Nash

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

		<u>Player 2</u>		
		Rock	Paper	Scissors
<u>Player 1</u>	Rock	0, 0	-1, 1	1, -1
	Paper	1, -1	0, 0	-1, 1
	Scissors	-1, 1	1, -1	0, 0

2-player normal-form game with finite set of actions taken simultaneously  
represented in a (bi)matrix

Player 1 is row player (typically first number) 

Player 2 is column player (typically second number) 

# 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>



# Strategy: Rock, Paper, Scissors

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

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

```
function playRPS(T)
```

```
Input:  $T$ 
```

```
Output: action  $a \in \{\text{Rock, Paper, Scissors}\}$ 
```

```
Return
```

*rock*

# 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 😞

Fill in the payoff matrix, row payoff first then column!

		Berry	
		Football	Concert
Alex	Football	<u>2</u> , 1	-1, <u>-1</u>
	Concert	-1, -1	1, 2

# Normal-Form Games

A game in normal form consists of the following elements

- Set of players  $Z$
- 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

$$U_A(a_A, a_B)$$

Players move simultaneously and the game ends immediately afterwards

What are the players, set of actions and utility functions of Football vs Concert (FvsC) game?

# Classical Games and Payoff Matrices

## Prisoner's Dilemma (PD)

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

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

## Variation: Split or Steal



<https://youtu.be/p3Uos2fzIJ0>



<https://www.wnycstudios.org/podcasts/radiolab/segments/golden-rule>

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

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

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

# Strategy

Pure strategy: choose an action deterministically

Mixed strategy: choose actions according to a probability distribution

- Notation:  $s = (\underline{0.3}, \underline{0.7}, \underline{0})$
- 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?

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

# Expected Utility

Given the strategies of all players,

Expected Utility for player  $i$   $u_i =$

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

Can skip action profiles with probability 0 or utility 0

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

$$[p_1, p_2, \dots, p_n]$$

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

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  $\mathbf{a} = (\text{Concert}, \text{Football})$ ?

Berry

What is Alex's utility in this action profile?

$$u_A = 1 + 0 + 0 + 0 = 1$$

	Football	Concert
Alex Football	$\frac{1}{2}$ 2,1	0 0,0
Concert	$\frac{1}{2}$ 0,0	0 1,2



# Poll 1

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. 7
- C. 6
- D. 4
- E. 3

		Player 2		
		0	$\frac{1}{2}$	$\frac{1}{2}$
		Rock	Paper	Scissors
Player 1	$\frac{1}{3}$	<del>0, 0</del>	-1, 1	1, -1
	$\frac{2}{3}$	1, -1	0, 0	-1, 1
	0	<del>-1, 1</del>	<del>1, -1</del>	<del>0, 0</del>

## Poll 2

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

$$\begin{aligned} & \cancel{\frac{1}{6}(-1)} + \cancel{\frac{1}{6}(1)} + \\ & \cancel{\frac{2}{6}(0)} + \frac{2}{6}(-1) \\ & = -\frac{1}{3} \end{aligned}$$

		Player 2		
		0	1/2	1/2
		Rock	Paper	Scissors
Player 1	1/3	<del>0, 0</del>	<u>1/6</u> -1, 1	<u>1/6</u> 1, -1
	2/3	<del>1, -1</del>	<u>2/6</u> 0, 0	<u>2/6</u> -1, 1
	0	<del>-1, 1</del>	<del>1, -1</del>	<del>0, 0</del>

# 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)?

In Prisoner's Dilemma, what is Player 1's best response to Cooperate? What is Player 1's best response to 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  $s_B = \left(\frac{1}{2}, \frac{1}{2}\right)$ ?

$$(P/2)2 + P/2(0) + (\cancel{1/2 - P/2})(0) + (\cancel{1/2 - P/2})1$$

$$P + 1/2 - P/2 = \boxed{1/2 + P}$$

		Berry	
		$\frac{1}{2}$ Football	$\frac{1}{2}$ Concert
Alex	Football $P$	$\frac{P}{2} \circledast 2, 1$	$\frac{P}{2} \circledast 0, 0$
	Concert $1-p$	$\frac{1}{2} - \frac{P}{2} \circledast 0, 0$	$\frac{1}{2} - \frac{P}{2} \circledast 1, 2$

## 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.  $s_2 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$  ← ○

$\frac{1}{3}$   
 $\frac{1}{3}$   
 $\frac{1}{3}$   
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

Player 2

# Best Response

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

		Player 2		
		Rock	Paper	Scissors
Player 1	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

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

strictly dominated  $\leftarrow$

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

$\leq$



$s_i$  **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

$\leq$



# Dominance

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

		Player 2		
		Rock	Paper	Scissors
Player 1	Rock	0,0	-1,1	1,-1
	Paper	1,-1	0,0	-1,1
	Scissor	-1,1	1,-1	0,0

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	<u>-1,-1</u>	<u>-3,0</u>
	Defect	<u>0,-3</u>	<u>-2,-2</u>

		Berry	
		Football	Concert
Alex	Football	<u>2,1</u>	<u>0,0</u>
	Concert	<u>0,0</u>	<u>1,2</u>

# Dominance

If  $s_i$  strictly dominates  $s'_i$ ,  $\forall s'_i \in S_i \setminus \{s_i\}$ ,  
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:

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

So... for any  $\mathbf{s}_{-i}$

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

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?

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

Note: Doesn't always exist

What is the dominant strategy equilibrium for PD?

		Player 2	
		Cooperate	Defect
Player 1	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)

- $a_i \in BR(\mathbf{a}_{-i}), \forall i$

## Mixed Strategy Nash Equilibrium

- At least one player use a randomized strategy
- $s_i \in BR(\mathbf{s}_{-i}), \forall i$

# Nash Equilibrium

What are the PSNEs in these games?

*pure*

What is the mixed strategy NE in RPS?

		Player 2		
		Rock	Paper	Scissors
Player 1	Rock	0,0	-1,1	1,-1
	Paper	1,-1	0,0	-1,1
	Scissor	-1,1	1,-1	0,0

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

		Berry	
		Football	Concert
Alex	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 profiles
- 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?

		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 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

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

Player 1

# Find Mixed Strategy Nash Equilibrium

Big idea:

A NE occurs when there's no incentive to change actions

Ensure that the expected utility of other player's actions is equal

Can we still apply iterative removal?

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

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

# 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  $q$ ?

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

$$u_B(s_A, F) = u_B(s_A, C)$$

# 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  $q$ ?

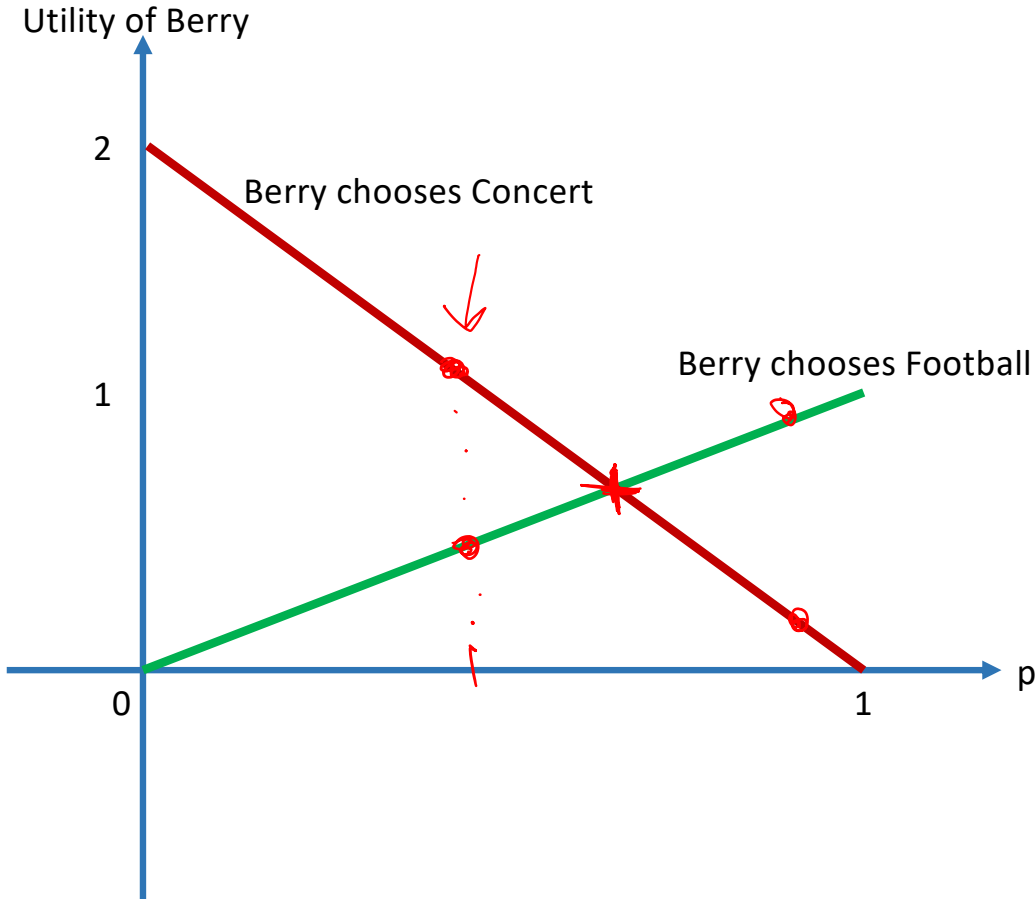
$$u_A(F, s_B) = u_A(C, s_B) \qquad u_B(s_A, F) = u_B(s_A, C)$$

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  $F \in BR(s_B)$  and  $C \in BR(s_B)$



# Game Theory as Linear Programs



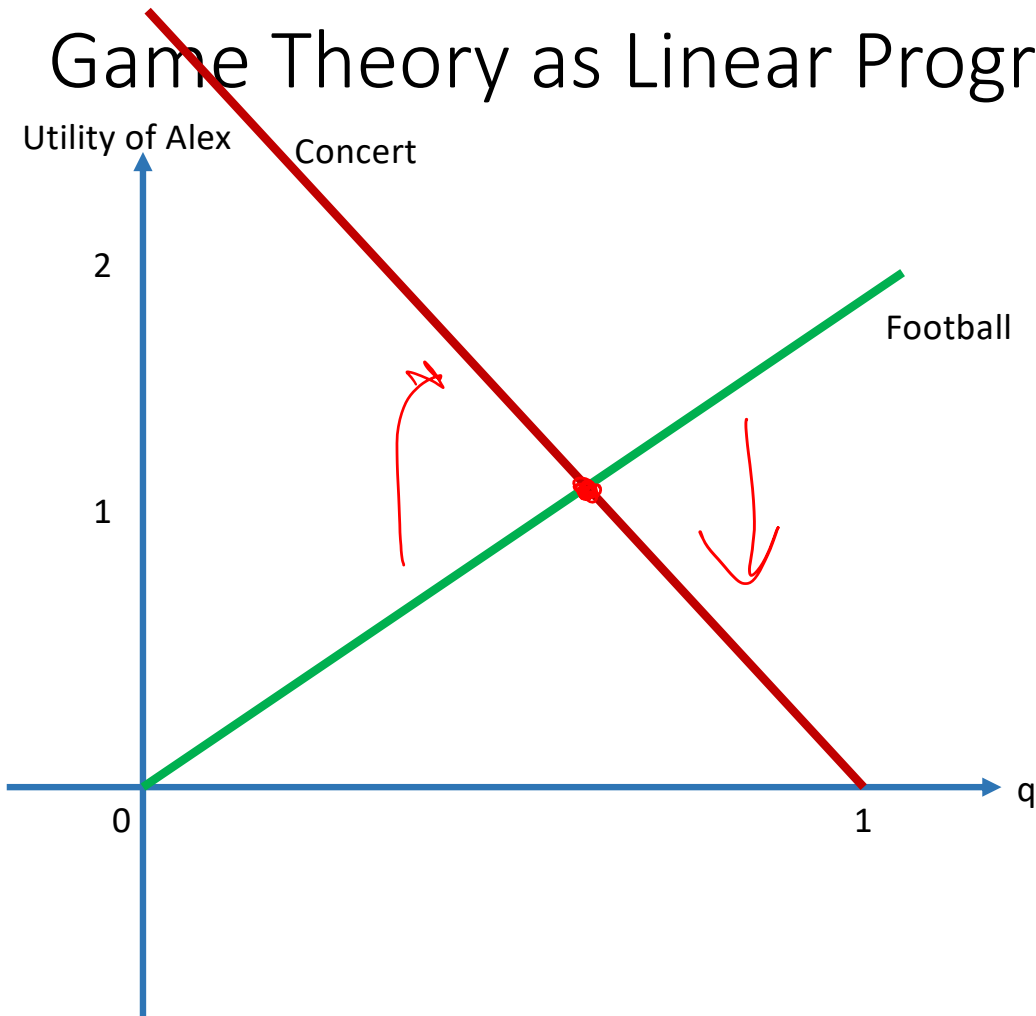
Alex wants to choose  $p$  such that Berry doesn't want to deviate from his strategy

Berry wants the most reward he can get, so he will deviate if one strategy has more utility than another

The only way these two conditions is met is if we choose the  $p$  such that any strategy Berry picks will yield equal utility for Berry

		Berry		
		Football	Concert	
Alex	$p$	Football	2,1	0,0
	$1-p$	Concert	0,0	3,2

# Game Theory as Linear Programs



Berry wants to choose  $q$  such that Alex doesn't want to deviate from his strategy

Alex wants the most reward he can get, so he will deviate if one strategy has more utility than another

The only way these two conditions is met is if we choose the  $q$  such that any strategy Alex picks will yield equal utility for Alex

		Berry	
		Football	Concert
Alex	Football	2,1	0,0
	Concert	0,0	3,2

## 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.  $2q = 3(1 - q)$
- 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	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?

What action should player 2 choose if player 1 commits to playing  $b$ ?

What is player 1's utility?

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

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

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

		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

# In-Class Activity

What is the Mixed Strategy Nash Equilibrium for this new problem?

		Berry	
		Football	Concert
Alex	Football	4,1	0,0
	Concert	0,0	3,3