

# Simple Games

**Next Year:** See notes within

- Last time:
- Today:

# Simple Games

**Next Year:** See notes within

- **Last time:**
  - Inference in Bayesian networks
  - Learning Bayesian networks
- **Today:**

# Simple Games

**Next Year:** See notes within

- **Last time:**
  - Inference in Bayesian networks
  - Learning Bayesian networks
- **Today:**
  - Games: a mathematical formalism for rational interaction
  - Nash and other equilibria

# Types of Uncertainty

# Types of Uncertainty

**Alleatory**

# Types of Uncertainty

**Alleatory**



**Markov Decision Process**

# Types of Uncertainty

**Alleatory**

**Epistemic (Static)**

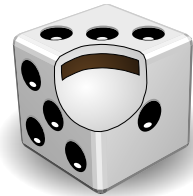


**Markov Decision Process**

# Types of Uncertainty

**Alleatory**

**Epistemic (Static)**



**Markov Decision Process**

**Reinforcement Learning**



# Types of Uncertainty

**Alleatory**



**Markov Decision Process**

**Epistemic (Static)**



**Reinforcement Learning**

**Epistemic (Dynamic)**



**POMDP**

# Types of Uncertainty

**Alleatory**



**Markov Decision Process**

**Epistemic (Static)**



**Reinforcement Learning**

**Epistemic (Dynamic)**



**POMDP**

**Interaction**

# Types of Uncertainty

**Alleatory**



**Markov Decision Process**

**Epistemic (Static)**



**Reinforcement Learning**

**Epistemic (Dynamic)**



**POMDP**

**Interaction**



**Game**

# A win-win situation: International trade

**Next Year:** This example isn't great

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: -1

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: -1
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: -1
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3
- No production capacity limits



# A win-win situation: International trade

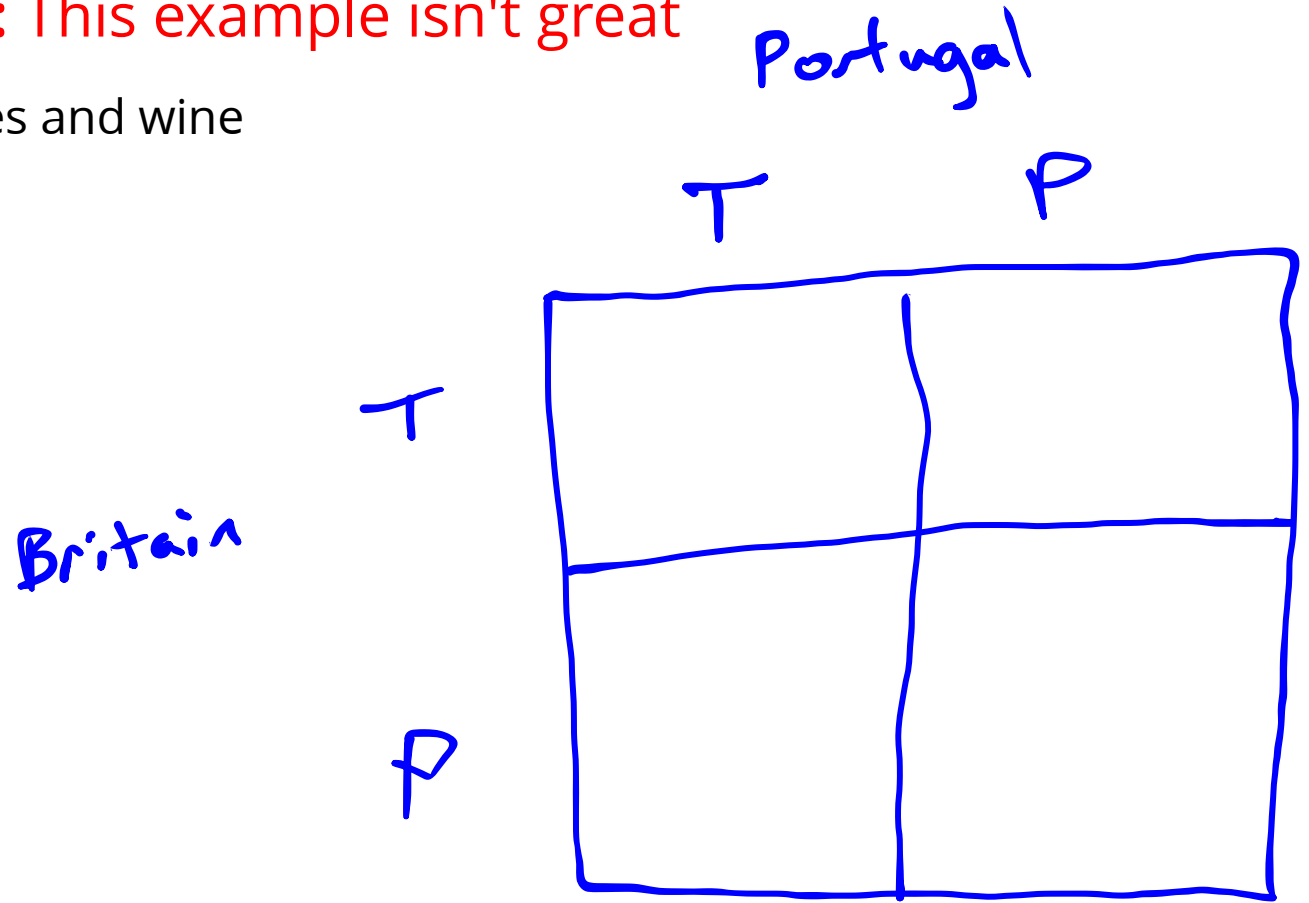
**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: -1
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3
- No production capacity limits
- Each country can either
  - Produce their own goods
  - Trade at a price of 2

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: -1
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3
- No production capacity limits
- Each country can either
  - Produce their own goods
  - Trade at a price of 2



# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: -1
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3
- No production capacity limits
- Each country can either
  - Produce their own goods
  - Trade at a price of 2

Portugal

T P

Britain

T P

	T	P
T		
P		-4, -4

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: -1
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3
- No production capacity limits
- Each country can either
  - Produce their own goods
  - Trade at a price of 2

Portugal

T P

Britain

T P

	T	P
T		-3, -3
P	-3, -3	-4, -4

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: -1
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3
- No production capacity limits
- Each country can either
  - Produce their own goods
  - Trade at a price of 2

Portugal

T P

Britain

T P

T	-2, -2	-3, -3
P	-3, -3	-4, -4

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: ~~-1~~ **-2**
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3
- No production capacity limits
- Each country can either
  - Produce their own goods
  - Trade at a price of 2

Britain

Portugal

		Portugal	
		T	P
Britain	T	-2, -2	-3, -3
	P	-3, -3	-4, -4

# A win-win situation: International trade

**Next Year:** This example isn't great

- Both Britain and Portugal need textiles and wine
- Britain:
  - Producing wine: -3
  - Producing textiles: ~~-1~~ **-2**
- Portugal:
  - Producing wine: -1
  - Producing textiles: -3
- No production capacity limits
- Each country can either
  - Produce their own goods
  - Trade at a price of 2

Britain

-4, -2	-4, -3
-5, -3	-5, -4

Portugal

	T	P
T	-2, -2	-3, -3
P	-3, -3	-4, -4

# A more surprising example: The Prisoner's Dilemma



# A more surprising example: The Prisoner's Dilemma

- 2 Alleged criminals are captured

# A more surprising example:

## The Prisoner's Dilemma

- 2 Alleged criminals are captured
- Each can either keep silent or testify
  - other keeps silent -> minor conviction (1 year)
  - other testifies -> major conviction: 4 years
  - testify -> 1 year removed from sentence

# Vocabulary and Notation

.

# Vocabulary and Notation

Single Player

Joint

# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$

# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_{-i}^i)$	$\pi(a) = \prod_i \pi^i(a_{-i}^i)$

# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$

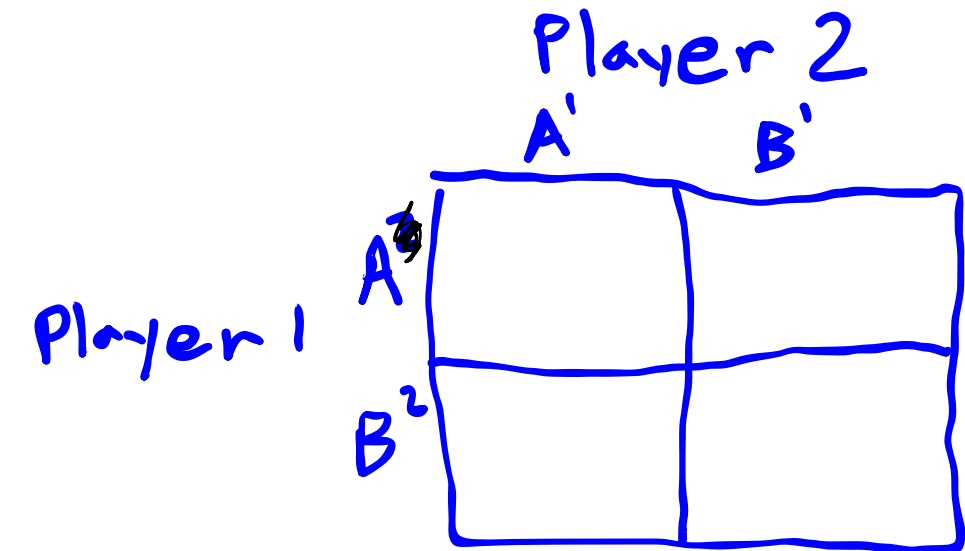
# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a)\pi(a)$	$U(\pi) = \sum_a R(a)\pi(a)$



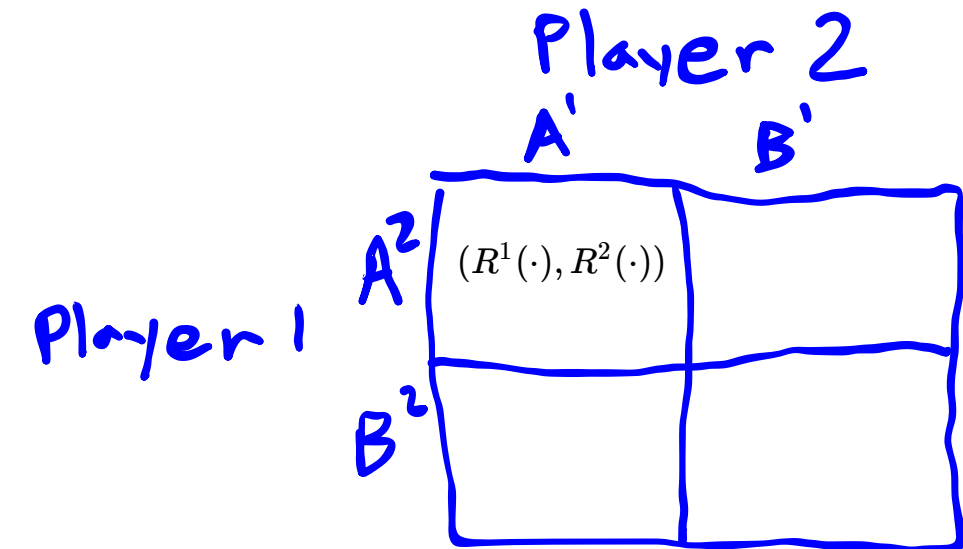
# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a)\pi(a)$	$U(\pi) = \sum_a R(a)\pi(a)$



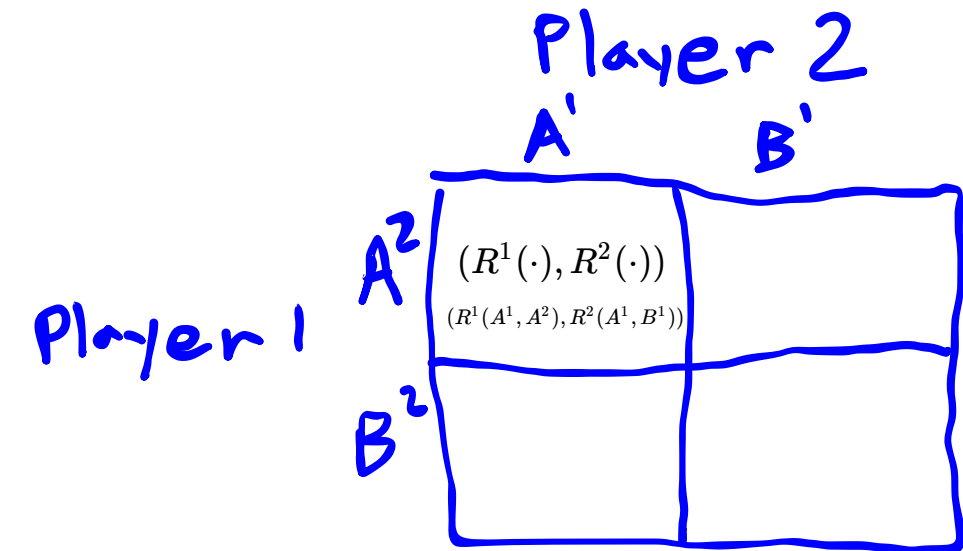
# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a)\pi(a)$	$U(\pi) = \sum_a R(a)\pi(a)$



# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a) \pi(a)$	$U(\pi) = \sum_a R(a) \pi(a)$



# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a) \pi(a)$	$U(\pi) = \sum_a R(a) \pi(a)$

Player 1

	Player 2	
	A'	B'
A <sup>2</sup>	$(R^1(\cdot), R^2(\cdot))$ $(R^1(A^1, A^2), R^2(A^1, B^1))$	$(R^1(\cdot), R^2(\cdot))$
B <sup>2</sup>		

# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a)\pi(a)$	$U(\pi) = \sum_a R(a)\pi(a)$

Handwritten game matrix:

		Player 2	
		A'	B'
Player 1	A <sup>2</sup>	$(R^1(\cdot), R^2(\cdot))$ $(R^1(A^1, A^2), R^2(A^1, B^1))$	$(R^1(\cdot), R^2(\cdot))$
	B <sup>2</sup>	$(R^1(\cdot), R^2(\cdot))$	

# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a)\pi(a)$	$U(\pi) = \sum_a R(a)\pi(a)$

Handwritten game matrix:

		Player 2	
		A'	B'
Player 1	A <sup>2</sup>	$(R^1(\cdot), R^2(\cdot))$ $(R^1(A^1, A^2), R^2(A^1, B^1))$	$(R^1(\cdot), R^2(\cdot))$
	B <sup>2</sup>	$(R^1(\cdot), R^2(\cdot))$	$(R^1(\cdot), R^2(\cdot))$

# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a) \pi(a)$	$U(\pi) = \sum_a R(a) \pi(a)$

Player 1

	Player 2	
	A'	B'
A <sup>2</sup>	$(R^1(\cdot), R^2(\cdot))$ $(R^1(A^1, A^2), R^2(A^1, B^1))$	$(R^1(\cdot), R^2(\cdot))$
B <sup>2</sup>	$(R^1(\cdot), R^2(\cdot))$	$(R^1(\cdot), R^2(\cdot))$

Robber 1

Robber 2


# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a)\pi(a)$	$U(\pi) = \sum_a R(a)\pi(a)$

Player 1

		Player 2	
		A'	B'
A <sup>2</sup>	B <sup>2</sup>	$(R^1(\cdot), R^2(\cdot))$ $(R^1(A^1, A^2), R^2(A^1, B^1))$	$(R^1(\cdot), R^2(\cdot))$
		$(R^1(\cdot), R^2(\cdot))$	$(R^1(\cdot), R^2(\cdot))$

Robber 1

		Robber 2	
		S	T
S	T		



# Vocabulary and Notation

	Single Player	Joint
• Action	$a^i \in A^i$	$a \in A$
• Policy (strategy)	$\pi^i(a_i)$	$\pi(a) = \prod_i \pi^i(a_i)$
• Reward	$R^i(a)$	$R(a)$
• Utility	$U^i(\pi) = \sum_a R^i(a)\pi(a)$	$U(\pi) = \sum_a R(a)\pi(a)$

$(R^1, R^2)$

Player 1

Player 2

	A	B
A	$(R^1(\cdot), R^2(\cdot))$ $(R^1(A^1, A^2), R^2(A^1, A^2))$	$(R^1(\cdot), R^2(\cdot))$
B	$(R^1(\cdot), R^2(\cdot))$	$(R^1(\cdot), R^2(\cdot))$

Robber 1

Robber 2

	S	T
S	-1, -1	-4, 0
T	0, -4	-3, -3

# Best Response

# Best Response

Let  $\pi^{-i}$  be the joint policy of all *other* players.

$$\pi^{-i} = (\underbrace{\pi^1, \dots, \pi^{i-1}, \pi^{i+1}, \dots, \pi^n}_{\pi^i})$$

# Best Response

Let  $\pi^{-i}$  be the joint policy of all *other* players.

$$\pi^{-i} = (\pi^1, \dots, \pi^{i-1}, \pi^{i+1}, \dots, \pi^n)$$

**Best Response:** Given a joint policy of all other agents,  $\pi^{-i}$ , a best response is a policy  $\pi^i$  that satisfies

$$U^i(\pi^i, \pi^{-i}) \geq U^i(\pi^{i'}, \pi^{-i})$$

for all other  $\pi^{i'}$ .

# Best Response

Let  $\pi^{-i}$  be the joint policy of all *other* players.

$$\pi^{-i} = (\pi^1, \dots, \pi^{i-1}, \pi^{i+1}, \dots, \pi^n)$$

**Best Response:** Given a joint policy of all other agents,  $\pi^{-i}$ , a best response is a policy  $\pi^i$  that satisfies

$$U^i(\pi^i, \pi^{-i}) \geq U^i(\pi^{i'}, \pi^{-i})$$

for all other  $\pi^{i'}$ .

Handwritten diagram illustrating a best response calculation for Robber 1.

At the top, two probabilities are given:

- $\pi^2(S) = 1.0$  (green text, green arrow pointing down to Robber 2's S column)
- $\pi^2(T) = 1.0$  (red text, red arrow pointing down to Robber 2's T column)

The game matrix is shown with Robber 2's strategies (S, T) as columns and Robber 1's strategies (S, T) as rows. The payoffs are (Robber 1, Robber 2):

	Robber 2 S	Robber 2 T
Robber 1 S	-1, -1	-4, 0
Robber 1 T	0, -4	-3, -3

Arrows indicate the best response for Robber 1:

- Green arrow pointing down from -1 to 0 in the first column (S vs T).
- Red arrow pointing down from 0 to -3 in the second column (S vs T).

Best Response for Robber 1  
Given Robber 2's strategy

# Dominant Strategy Equilibria

# Dominant Strategy Equilibria

## Dominant Strategy Equilibrium

# Dominant Strategy Equilibria

## Dominant Strategy Equilibrium

- A *dominant strategy* is a policy that is a best response to all other possible agent policies.



# Dominant Strategy Equilibria

## Dominant Strategy Equilibrium

- A *dominant strategy* is a policy that is a best response to all other possible agent policies.

Robber 2

	S	T
Robber 1 S	-1, -1	-4, 0
T	0, -4	-3, -3

Handwritten annotations: Green arrows point from the top row (Robber 1's S and T) to the bottom row (Robber 1's T). Red arrows point from the left column (Robber 2's S and T) to the right column (Robber 2's T).

Testifying is a  
dominant strategy for  
Robber 1  
Robber 2

# Dominant Strategy Equilibria

## Dominant Strategy Equilibrium

- A *dominant strategy* is a policy that is a best response to all other possible agent policies.
- A joint policy where all agents use a dominant strategy is called a *dominant strategy equilibrium*.

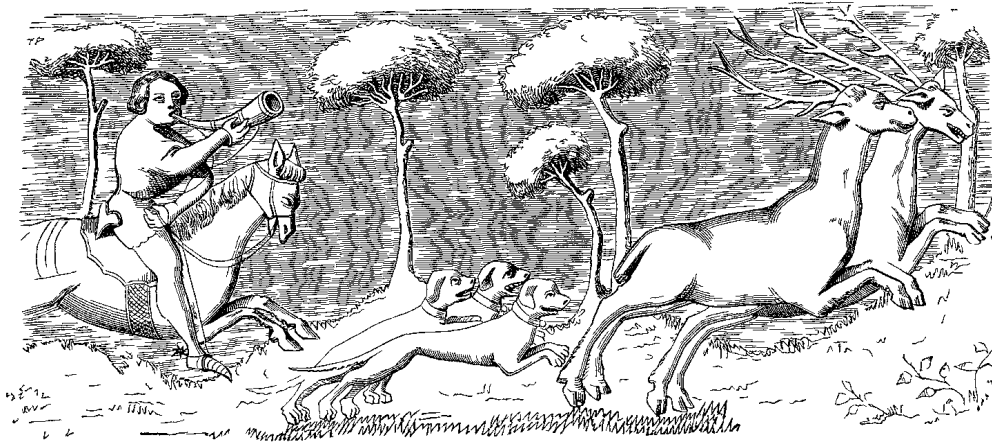
Robber 2

	S	T
Robber 1 S	-1, -1	-4, 0
T	0, -4	-3, -3

Both Robbers  
Testifying is  
a Dominant  
Strategy  
Equilibrium

# Nash Equilibria

# Nash Equilibria



	Stag	Hare
Stag	4, 4	1, 3
Hare	3, 1	2, 2

# Nash Equilibria

## Nash Equilibrium

	Stag	Hare
Stag	4, 4	1, 3
Hare	3, 1	2, 2

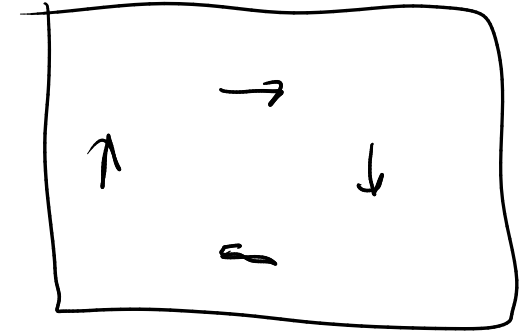
# Nash Equilibria

## Nash Equilibrium

- A *Nash equilibrium* is a joint policy in which all agents are following a best response

Identifying Pure Nash Equilibria

1. Calculate Best responses of all players (draw arrows)
2. if a square doesn't have any out arrows it is a NE



Player 2

	Stag	Hare
Stag	4, 4	1, 3
Hare	3, 1	2, 2

Player 1

# Geopolitics

# Rock-paper scissors

$\pi^2 = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$   
 is  $\pi^*$  a best response?  
 $U(\pi^*, \pi^2) \geq U(\pi'', \pi^2)$   
 No  
 $\pi'' = (0, 0, 1)$

		agent 2		
		rock	paper	scissors
agent 1	rock	0, 0	-1, 1	1, -1
	paper	1, -1	0, 0	-1, 1
	scissors	-1, 1	1, -1	0, 0

A two-player game is **zero sum** if

$$\sum_i R^i(a) = 0 \quad \forall a$$

- Pure strategy:  $\pi^i(a^i) \in \{0, 1\}$
- Mixed strategy: all other strategies

since two-player  
 $\pi^1 = \pi^2$

Consider the strategy  $\pi^*$  where

$$(\pi^{*,i}(a^i) = \frac{1}{3} \quad \forall i, a^i)$$

Prove that this is a NE

Definition of Best Response

$$U'(\pi^*, \pi^2) \geq U'(\pi'', \pi^2) \quad \forall \pi''$$

$$U'(\pi^*, \pi^2) = \sum_a R^1(a) \pi^*(a) = \frac{1}{3} \cdot \frac{1}{3} (0 + -1 + 1 + 1 + 0 + -1 + -1 + 1 + 0) = 0$$

$$U'(\pi'', \pi^2) = \sum_a R^1(a) \pi''(a) \pi^2(a) = 0 \cdot p_r \cdot \frac{1}{3} + -1 \cdot p_r \cdot \frac{1}{3} + 1 \cdot p_r \cdot \frac{1}{3} + \dots + \dots$$

$$U'(\pi'', \pi^2) = 0^a$$

$$\pi'' = (p_r, p_p, 1 - p_r - p_p)$$



# General approach to find Nash Equilibria

# General approach to find Nash Equilibria

if 0 then  $\pi$  is NE

$$\begin{aligned} &\underset{\pi, U}{\text{minimize}} && \sum_i \left( \underline{U^i} - U^i(\pi) \right) \\ &\text{subject to} && U^i \geq \underline{U^i(a^i, \pi^{-i})} \text{ for all } i, a^i \\ &&& \sum_{a^i} \pi^i(a^i) = 1 \text{ for all } i \\ &&& \pi^i(a^i) \geq 0 \text{ for all } i, a^i \end{aligned}$$

# General approach to find Nash Equilibria

Any number of players, any number of actions

$$\text{minimize}_{\pi, U} \sum_i (U^i - U^i(\pi))$$

$$\text{subject to } U^i \geq U^i(a^i, \pi^{-i}) \text{ for all } i, a^i$$

$$\sum_{a^i} \pi^i(a^i) = 1 \text{ for all } i$$

$$\pi^i(a^i) \geq 0 \text{ for all } i, a^i$$

root-finding problem

4	1	6	5	4	3	2	1	6	5	4	3	2	1
2	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Heg. Stability	Samaritan <sub>100</sub>	Samaritan <sub>100</sub>	Clock <sub>100</sub>	Clock <sub>100</sub>	Endless	Called Bluff	Bully	Unfair	Skewed BoS	Asym BoS	Chicken		
3	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Samson	Asym Sd <sub>100</sub>	Asym Sd <sub>100</sub>	Cycle <sub>100</sub>	Cycle <sub>100</sub>	Inspector	Self-serving <sub>100</sub>	Protector <sub>100</sub>	Protector <sub>100</sub>	Favorites <sub>100</sub>	Battle of Sexes	Asym BoS		
4	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Delliah	Asym Sd <sub>100</sub>	Asym Sd <sub>100</sub>	Pursuit	Pareto	Missile Crisis	Self-serving <sub>100</sub>	Protector <sub>100</sub>	Protector <sub>100</sub>	Hero	Favorites <sub>100</sub>	Skewed BoS		
5	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Hostage	Benevolence <sub>100</sub>	Benevolence <sub>100</sub>	2nd Best <sub>100</sub>	2nd Best <sub>100</sub>	Big Bully	Tragedy	Delight <sub>100</sub>	Pure Delight	Protector <sub>100</sub>	Protector <sub>100</sub>	Unfair		
6	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Blackmailer	Benevolence <sub>100</sub>	Benevolence <sub>100</sub>	2nd Best <sub>100</sub>	2nd Best <sub>100</sub>	Hamlet	Total Conflict	Mixed Delight	Delight <sub>100</sub>	Protector <sub>100</sub>	Protector <sub>100</sub>	Bully		
1	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Id. Hegemony	Samaritan <sub>100</sub>	Samaritan <sub>100</sub>	Revelation	Alibi	Asym Pd	Prisoners D.	Total Conflict	Tragedy	Self-serving <sub>100</sub>	Self-serving <sub>100</sub>	Called Bluff		
2	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Win-win	C. Aligned <sub>100</sub>	C. Aligned <sub>100</sub>	C. Assurance <sub>100</sub>	C. Assurance <sub>100</sub>	Stag Hunt	Asym Pd	Hamlet	Big Bully	Missile Crisis	Inspector	Endless		
3	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
R Assurance	Commons <sub>100</sub>	Commons <sub>100</sub>	Coordination <sub>100</sub>	Coordination <sub>100</sub>	R Assurance	Alibi	2nd Best <sub>100</sub>	2nd Best <sub>100</sub>	Pareto	Cycle <sub>100</sub>	Clock <sub>100</sub>		
4	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
R Assurance	Commons <sub>100</sub>	Commons <sub>100</sub>	Coordination <sub>100</sub>	Coordination <sub>100</sub>	R Assurance	Revelation	2nd Best <sub>100</sub>	2nd Best <sub>100</sub>	Pursuit	Cycle <sub>100</sub>	Clock <sub>100</sub>		
5	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Row Aligned	Harmony <sub>100</sub>	Harmony-mix	Commons <sub>100</sub>	Commons <sub>100</sub>	Row Aligned <sub>100</sub>	Samaritan <sub>100</sub>	Benevolent <sub>100</sub>	Benevolent <sub>100</sub>	Asym Sd <sub>100</sub>	Asym Sd <sub>100</sub>	Samaritan <sub>100</sub>		
6	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
Row Aligned	Harmony-pure	Harmony	Commons <sub>100</sub>	Commons <sub>100</sub>	Row Aligned <sub>100</sub>	Samaritan <sub>100</sub>	Benevolent <sub>100</sub>	Benevolent <sub>100</sub>	Asym Sd <sub>100</sub>	Asym Sd <sub>100</sub>	ActiveSam <sub>100</sub>		
1	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4
No Conflict	C. Aligned <sub>100</sub>	C. Aligned <sub>100</sub>	C. Assurance <sub>100</sub>	C. Assurance <sub>100</sub>	Win-win	Id. Hegemony	Blackmailer	Hostage	Delliah	Samson	Heg. Stability		
3	3	4	2	3	4	2	3	4	2	3	4	2	3
1	4	1	4	1	4	1	4	1	4	1	4	1	4

**Every finite game has a Nash Equilibrium**

# Every finite game has a Nash Equilibrium

## *EQUILIBRIUM POINTS IN $N$ -PERSON GAMES*

BY JOHN F. NASH, JR.\*

PRINCETON UNIVERSITY

Communicated by S. Lefschetz, November 16, 1949

One may define a concept of an  $n$ -person game in which each player has a finite set of pure strategies and in which a definite set of payments to the  $n$  players corresponds to each  $n$ -tuple of pure strategies, one strategy being taken for each player. For mixed strategies, which are probability distributions over the pure strategies, the pay-off functions are the expectations of the players, thus becoming polylinear forms in the probabilities with which the various players play their various pure strategies.

Any  $n$ -tuple of strategies, one for each player, may be regarded as a point in the product space obtained by multiplying the  $n$  strategy spaces of the players. One such  $n$ -tuple counters another if the strategy of each player in the countering  $n$ -tuple yields the highest obtainable expectation for its player against the  $n - 1$  strategies of the other players in the countered  $n$ -tuple. A self-countering  $n$ -tuple is called an equilibrium point.

The correspondence of each  $n$ -tuple with its set of countering  $n$ -tuples gives a one-to-many mapping of the product space into itself. From the definition of countering we see that the set of countering points of a point is convex. By using the continuity of the pay-off functions we see that the graph of the mapping is closed. The closedness is equivalent to saying: if  $P_1, P_2, \dots$  and  $Q_1, Q_2, \dots, Q_n, \dots$  are sequences of points in the product space where  $Q_n \rightarrow Q$ ,  $P_n \rightarrow P$  and  $Q_n$  counters  $P_n$  then  $Q$  counters  $P$ .

Since the graph is closed and since the image of each point under the mapping is convex, we infer from Kakutani's theorem<sup>1</sup> that the mapping has a fixed point (i.e., point contained in its image). Hence there is an equilibrium point.

In the two-person zero-sum case the "main theorem"<sup>2</sup> and the existence of an equilibrium point are equivalent. In this case any two equilibrium points lead to the same expectations for the players, but this need not occur in general.

\* The author is indebted to Dr. David Gale for suggesting the use of Kakutani's theorem to simplify the proof and to the A. E. C. for financial support.


<sup>1</sup> Kakutani, S., *Duke Math. J.*, 8, 457-459 (1941).

<sup>2</sup> Von Neumann, J., and Morgenstern, O., *The Theory of Games and Economic Behaviour*, Chap. 3, Princeton University Press, Princeton, 1947.

# Every finite game has a Nash Equilibrium

## **Kakutani's fixed-point theorem**

A correspondence  $f: X \rightarrow X$  has a fixed point (i.e.,  $\mathbf{x} \in f(\mathbf{x})$  for some  $\mathbf{x} \in X$ ) if all of the following conditions hold.

- (1)  $X$  is a non-empty, closed, bounded, and convex set.
  - (2)  $f(\mathbf{x})$  is non-empty for any  $\mathbf{x}$ .
  - (3)  $f(\mathbf{x})$  is convex for any  $\mathbf{x}$ .
  - (4) The set  $\{ (\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in f(\mathbf{x}) \}$  is closed.
- 

# Every finite game has a Nash Equilibrium

## **Kakutani's fixed-point theorem**

A correspondence  $f: X \rightarrow X$  has a fixed point (i.e.,  $\mathbf{x} \in f(\mathbf{x})$  for some  $\mathbf{x} \in X$ ) if all of the following conditions hold.

- (1)  $X$  is a non-empty, closed, bounded, and convex set.
- (2)  $f(\mathbf{x})$  is non-empty for any  $\mathbf{x}$ .
- (3)  $f(\mathbf{x})$  is convex for any  $\mathbf{x}$ .
- (4) The set  $\{ (\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in f(\mathbf{x}) \}$  is closed.

- Let  $x$  be a <sup>Joint policy</sup> strategy profile,  $\pi$ .

# Every finite game has a Nash Equilibrium

## **Kakutani's fixed-point theorem**

A correspondence  $f: X \rightarrow X$  has a fixed point (i.e.,  $\mathbf{x} \in f(\mathbf{x})$  for some  $\mathbf{x} \in X$ ) if all of the following conditions hold.

- (1)  $X$  is a non-empty, closed, bounded, and convex set.
- (2)  $f(\mathbf{x})$  is non-empty for any  $\mathbf{x}$ .
- (3)  $f(\mathbf{x})$  is convex for any  $\mathbf{x}$ .
- (4) The set  $\{ (\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in f(\mathbf{x}) \}$  is closed.

- Let  $x$  be a strategy profile,  $\pi$ .
- Let  $f$  be  $BR_i$ , that is, the best response operator



# Every finite game has a Nash Equilibrium

## **Kakutani's fixed-point theorem**

A correspondence  $f: X \rightarrow X$  has a fixed point (i.e.,  $\mathbf{x} \in f(\mathbf{x})$  for some  $\mathbf{x} \in X$ ) if all of the following conditions hold.

- (1)  $X$  is a non-empty, closed, bounded, and convex set.
- (2)  $f(\mathbf{x})$  is non-empty for any  $\mathbf{x}$ .
- (3)  $f(\mathbf{x})$  is convex for any  $\mathbf{x}$ .
- (4) The set  $\{ (\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in f(\mathbf{x}) \}$  is closed.

- Let  $x$  be a strategy profile,  $\pi$ .
- Let  $f$  be  $BR$ , that is, the best response operator
- A fixed point of  $BR$  is a Nash Equilibrium

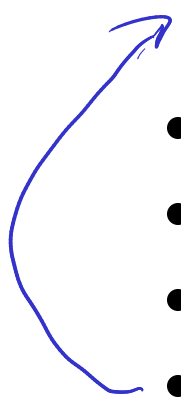
# Every finite game has a Nash Equilibrium

## Kakutani's fixed-point theorem

A correspondence  $f: X \rightarrow X$  has a fixed point (i.e.,  $\mathbf{x} \in f(\mathbf{x})$  for some  $\mathbf{x} \in X$ ) if all of the following conditions hold.

- (1)  $X$  is a non-empty, closed, bounded, and convex set.
- (2)  $f(\mathbf{x})$  is non-empty for any  $\mathbf{x}$ .
- (3)  $f(\mathbf{x})$  is convex for any  $\mathbf{x}$ .
- (4) The set  $\{ (\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in f(\mathbf{x}) \}$  is closed.

for VI  
Banach's Theorem

- 
- Let  $x$  be a strategy profile,  $\pi$ .
  - Let  $f$  be  $BR$ , that is, the best response operator
  - A fixed point of  $BR$  is a Nash Equilibrium
  - The  $BR$  operator and policy space for finite games meet the conditions above

# Every finite game has a Nash Equilibrium

## **Kakutani's fixed-point theorem**

A correspondence  $f: X \rightarrow X$  has a fixed point (i.e.,  $\mathbf{x} \in f(\mathbf{x})$  for some  $\mathbf{x} \in X$ ) if all of the following conditions hold.

- (1)  $X$  is a non-empty, closed, bounded, and convex set.
- (2)  $f(\mathbf{x})$  is non-empty for any  $\mathbf{x}$ .
- (3)  $f(\mathbf{x})$  is convex for any  $\mathbf{x}$ .
- (4) The set  $\{ (\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in f(\mathbf{x}) \}$  is closed.

- Let  $x$  be a strategy profile,  $\pi$ .
- Let  $f$  be  $BR$ , that is, the best response operator
- A fixed point of  $BR$  is a Nash Equilibrium
- The  $BR$  operator and policy space for finite games meet the conditions above
- $BR$  has a fixed point for every finite game, i.e. every finite game has a Nash Equilibrium

# Battle of the Sexes

# Battle of the Sexes

- Gabby and Max are going on a date

# Battle of the Sexes

- Gabby and Max are going on a date
- Gabby wants to go to a football game

# Battle of the Sexes

- Gabby and Max are going on a date
- Gabby wants to go to a football game
- Max wants to go to a movie (He is a rom-com superfan)

# Battle of the Sexes

- Gabby and Max are going on a date
- Gabby wants to go to a football game
- Max wants to go to a movie (He is a rom-com superfan)

A hand-drawn blue payoff matrix for the Battle of the Sexes game. The matrix is a square divided into four quadrants by a horizontal and vertical line. The columns are labeled 'G' and 'M' at the top, and the rows are labeled 'G' and 'M' on the left. The quadrants are empty, representing the payoffs for each combination of choices.

	G	M
G		
M		



# Battle of the Sexes

- Gabby and Max are going on a date
- Gabby wants to go to a football game
- Max wants to go to a movie (He is a rom-com superfan)

$\pi^G\left(\frac{2}{3}, \frac{1}{3}\right)$   
 $\pi^M\left(\frac{1}{3}, \frac{2}{3}\right)$

	M	
	G	M
G	2, 1	0, 0
	0, 2	1, 2

Handwritten notes on the table:

- A green box highlights the top-left cell (2, 1).
- A red arrow points from the top-right cell (0, 0) to the top-left cell (2, 1).
- A green arrow points from the bottom-left cell (0, 2) to the top-left cell (2, 1).
- A red arrow points from the bottom-left cell (0, 2) to the bottom-right cell (1, 2).
- A green arrow points from the bottom-right cell (1, 2) to the top-right cell (0, 0).

# Battle of the Sexes

- Gabby and Max are going on a date
- Gabby wants to go to a football game
- Max wants to go to a movie (He is a rom-com superfan)

Uncorrelated  
 $\pi(a) \neq \prod \pi^i(a^i)$

## Correlated Equilibrium

- A *correlated joint policy* is a single distribution over the joint actions of all agents.
- A *correlated equilibrium* is a correlated joint policy where no agent  $i$  can increase their expected utility by deviating from their current action to another.

Handwritten payoff matrix for the Battle of the Sexes game:

		M	
		G	M
G	G	2, 1	0, 0
	M	0, 2	1, 2

$$\begin{aligned}\pi(G, G) &= \frac{1}{2} \\ \pi(M, M) &= \frac{1}{2} \\ \pi(\cdot, \cdot) &= 0\end{aligned}$$

$$\begin{aligned}\pi(G, G) &= a \\ \pi(M, M) &= 1 - a \\ \pi(\cdot, \cdot) &= 0\end{aligned}$$

# Recap

# Recap

- Games provide a mathematical framework for analyzing interaction between rational agents

# Recap

- Games provide a mathematical framework for analyzing interaction between rational agents
- Games may not have a single "optimal" solution; instead there are equilibria

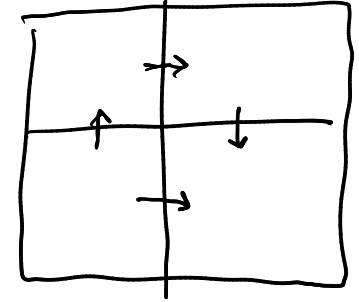
# Recap

- Games provide a mathematical framework for analyzing interaction between rational agents
- Games may not have a single "optimal" solution; instead there are equilibria
- If every player is playing a best response, that joint policy is a Nash Equilibrium

# Recap

- Games provide a mathematical framework for analyzing interaction between rational agents
- Games may not have a single "optimal" solution; instead there are equilibria
- If every player is playing a best response, that joint policy is a Nash Equilibrium
- Every finite game has at least one Nash Equilibrium (pure or mixed)

# Practice



Player 1

	Player 2		
	a	b	c
a	4,4	2,5	0,0
b	5,2	3,3	0,0
c	0,0	0,0	10,10