

# Game Theory

CMSC 421, Section 17.6

# Introduction

- In Chapter 6 we looked at 2-player perfect-information zero-sum games
- We'll now look at games that might have one or more of the following:
  - $> 2$  players
  - imperfect information
  - nonzero-sum outcomes

# The Prisoner's Dilemma

- Scenario: the police have arrested two suspects for a crime.
  - They tell each prisoner they'll reduce his/her prison sentence if he/she betrays the other prisoner.
  - Each prisoner must choose between two actions:
    - cooperate with the other prisoner, i.e., don't betray him/her
    - defect (betray the other prisoner).

- Payoff =  $-(\text{years in prison})$ :

- Each player has only two strategies, each of which is a single action

- Non-zero-sum

- Imperfect information: neither player knows the other's move until after *both* players have moved

Prisoner's Dilemma

Agent 2 \ Agent 1	C	D
C	-2, -2	-5, 0
D	0, -5	-4, -4

# The Prisoner's Dilemma

- Add 5 to each payoff, so that the numbers are all  $\geq 0$ 
  - These payoffs encode the same preferences

Prisoner's Dilemma:

Agent 1 \ Agent 2	C	D
C	-2, -2	-5, 0
D	0, -5	-4, -4

Prisoner's Dilemma:

Agent 1 \ Agent 2	C	D
C	3, 3	0, 5
D	5, 0	1, 1

- Note: the book represents payoff matrices in a non-standard way
  - It puts Agent 1 where I have Agent 2, and vice versa

# How to reason about games?

- In single-agent decision theory, look at an **optimal strategy**
  - Maximize the agent's expected payoff in its environment
- With multiple agents, the best strategy depends on others' choices
- Deal with this by identifying certain subsets of outcomes called **solution concepts**
- Some solution concepts:
  - Dominant strategy equilibrium
  - Pareto optimality
  - Nash equilibrium

# Strategies

- Suppose the agents agent 1, agent 2, ..., agent  $n$
- For each  $i$ , let  $S_i = \{\text{all possible strategies for agent } i\}$ 
  - $s_i$  will always refer to a strategy in  $S_i$
- A **strategy profile** is an  $n$ -tuple  $S = (s_1, \dots, s_n)$ , one strategy for each agent
- **Utility**  $U_i(S) = \text{payoff for agent } i \text{ if the strategy profile is } S$
- $s_i$  **strongly dominates**  $s_i'$  if agent  $i$  always does better with  $s_i$  than  $s_i'$

$$\forall s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n,$$

$$U_i(s_1, \dots, s_{i-1}, s_i, s_{i+1}, \dots, s_n) > U_i(s_1, \dots, s_{i-1}, s_i', s_{i+1}, \dots, s_n)$$

- $s_i$  **weakly dominates**  $s_i'$  if agent  $i$  never does worse with  $s_i$  than  $s_i'$ , and there is at least one case where agent  $i$  does better with  $s_i$  than  $s_i'$ ,

$$\forall s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n, U_i(\dots, s_i, \dots) \geq U_i(\dots, s_i', \dots)$$

and

$$\exists s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n \quad U_i(\dots, s_i, \dots) > U_i(\dots, s_i', \dots)$$

# Dominant Strategy Equilibrium

- $s_i$  is a **(strongly, weakly) dominant** strategy if it (strongly, weakly) dominates every  $s_i' \in S_i$
- **Dominant strategy equilibrium:**
  - A set of strategies  $(s_1, \dots, s_n)$  such that each  $s_i$  is dominant for agent  $i$
  - Thus agent  $i$  will do best by using  $s_i$  rather than a different strategy, regardless of what strategies the other players use
  - In the prisoner's dilemma, there is one dominant strategy equilibrium: both players defect

Prisoner's Dilemma:

	Agent 2	C	D
Agent 1			
C		3, 3	0, 5
D		5, 0	1, 1

# Pareto Optimality

- Strategy profile  $S$  **Pareto dominates** a strategy profile  $S'$  if
  - no agent gets a worse payoff with  $S$  than with  $S'$ ,  
i.e.,  $U_i(S) \geq U_i(S')$  for all  $i$ ,
  - at least one agent gets a better payoff with  $S$  than with  $S'$ ,  
i.e.,  $U_i(S) > U_i(S')$  for at least one  $i$
- Strategy profile  $s$  is **Pareto optimal**, or **strictly Pareto efficient**, if there's no strategy  $s'$  that Pareto dominates  $s$ 
  - Every game has at least one Pareto optimal profile
  - Always at least one Pareto optimal profile in which the strategies are pure

# Example

## The Prisoner's Dilemma

- $(C, C)$  is Pareto optimal
  - No profile gives both players a higher payoff
- $(D, C)$  is Pareto optimal
  - No profile gives player 1 a higher payoff
- $(D, C)$  is Pareto optimal - same argument
- $(D, D)$  is Pareto dominated by  $(C, C)$ 
  - But ironically,  $(D, D)$  is the dominant strategy equilibrium

Prisoner's Dilemma

Agent 1 \ Agent 2	C	D
C	3, 3	0, 5
D	5, 0	1, 1

# Pure and Mixed Strategies

- **Pure strategy:** select a single action and play it
  - Each row or column of a payoff matrix represents both an action and a pure strategy
- **Mixed strategy:** randomize over the set of available actions according to some probability distribution
  - Let  $A_i = \{\text{all possible actions for agent } i\}$ , and  $a_i$  be any action in  $A_i$
  - $s_i(a_j) = \text{probability that action } a_j \text{ will be played under mixed strategy } s_i$
- The **support** of  $s_i$  is
  - $\text{support}(s_i) = \{\text{actions in } A_i \text{ that have probability } > 0 \text{ under } s_i\}$
- A pure strategy is a special case of a mixed strategy
  - support consists of a single action
- **Fully mixed strategy:** every action has probability  $> 0$ 
  - i.e.,  $\text{support}(s_i) = A_i$

# Expected Utility

- A payoff matrix only gives payoffs for pure-strategy profiles
- Generalization to mixed strategies uses *expected utility*
- Let  $S = (s_1, \dots, s_n)$  be a profile of mixed strategies
  - For every action profile  $(a_1, a_2, \dots, a_n)$ , multiply its probability and its utility
    - $U_i(a_1, \dots, a_n) s_1(a_1) s_2(a_2) \dots s_n(a_n)$
  - The expected utility for agent  $i$  is

$$U_i(s_1, \dots, s_n) = \sum_{(a_1, \dots, a_n) \in \mathbf{A}} U_i(a_1, \dots, a_n) s_1(a_1) s_2(a_2) \dots s_n(a_n)$$

# Best Response

- Some notation:
  - If  $S = (s_1, \dots, s_n)$  is a strategy profile, then  $S_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ ,
    - i.e.,  $S_{-i}$  is strategy profile  $S$  without agent  $i$ 's strategy
  - If  $s_i'$  is any strategy for agent  $i$ , then
    - $(s_i', S_{-i}) = (s_1, \dots, s_{i-1}, s_i', s_{i+1}, \dots, s_n)$
  - Hence  $(s_i, S_{-i}) = S$
- $s_i$  is a **best response** to  $S_{-i}$  if
$$U_i(s_i, S_{-i}) \geq U_i(s_i', S_{-i})$$
 for every strategy  $s_i'$  available to agent  $i$
- $s_i$  is a **unique** best response to  $S_{-i}$  if
$$U_i(s_i, S_{-i}) > U_i(s_i', S_{-i})$$
 for every  $s_i' \neq s_i$

# Nash Equilibrium

- A strategy profile  $s = (s_1, \dots, s_n)$  is a **Nash equilibrium** if for every  $i$ ,
  - $s_i$  is a best response to  $S_{-i}$ , i.e., no agent can do better by unilaterally changing his/her strategy
- **Theorem (Nash, 1951):** Every game with a finite number of agents and action profiles has at least one Nash equilibrium
- In the Prisoner's Dilemma,  $(D,D)$  is a Nash equilibrium
  - If either agent unilaterally switches to a different strategy, his/her expected utility goes below 1
- A dominant strategy equilibrium is always a Nash equilibrium

Prisoner's Dilemma

	Agent 2	
Agent 1 \	C	D
C	3, 3	0, 5
D	5, 0	1, 1

# Example



- **Battle of the Sexes**

- Two agents need to coordinate their actions, but they have different preferences

- Original scenario:

- husband prefers football
- wife prefers opera

- Another scenario:

- Two nations must act together to deal with an international crisis
- They prefer different solutions

- This game has two pure-strategy Nash equilibria (circled above) and one mixed-strategy Nash equilibrium

- How to find the mixed-strategy Nash equilibrium?

	Husband	Opera	Football
Wife			
Opera		2, 1	0, 0
Football		0, 0	1, 2

Nash equilibria

# Finding Mixed-Strategy Equilibria

- Generally it's tricky to compute mixed-strategy equilibria
  - But easy if we can identify the support of the equilibrium strategies
- Suppose a best response to  $S_{-i}$  is a mixed strategy  $s$  whose support includes  $\geq 2$  actions
  - Then every action  $a$  in  $\text{support}(s)$  must have the same expected utility  $U_i(a, S_{-i})$ 
    - If some action  $a^*$  in  $\text{support}(s)$  had a higher expected utility than the others, then it would be a better response than  $s$
  - Thus *any* mixture of the actions in  $\text{support}(s)$  is a best response

# Battle of the Sexes

- Suppose both agents randomize, and the husband's mixed strategy  $s_h$  is

$$s_h(\text{Opera}) = p; \quad s_h(\text{Football}) = 1 - p$$

- Expected utilities of the wife's actions:

$$U_w(\text{Football}, s_h) = 0p + 1(1 - p)$$

$$U_w(\text{Opera}, s_h) = 2p$$

	Husband	
	Opera	Football
Wife	Opera	Football
	2, 1	0, 0
	0, 0	1, 2

- If the wife mixes between her two actions, they must have the same expected utility
  - If one of the actions had a better expected utility, she'd do better with a pure strategy that *always* used that action
  - Thus  $0p + 1(1 - p) = 2p$ , so  $p = 1/3$
- So the husband's mixed strategy is  $s_h(\text{Opera}) = 1/3$ ;  $s_h(\text{Football}) = 2/3$

# Battle of the Sexes

- A similar calculation shows that the wife's mixed strategy  $s_w$  is

$$s_w(\text{Opera}) = 2/3, \quad s_w(\text{Football}) = 1/3$$

- In this equilibrium,

- $P(\text{wife gets 2, husband gets 1})$   
 $= (2/3) (1/3) = 2/9$

- $P(\text{wife gets 1, husband gets 2})$   
 $= (1/3) (2/3) = 2/9$

- $P(\text{both get 0}) = (1/3)(1/3) + (2/3)(2/3) = 5/9$

- Thus the expected utility for each agent is  $2/3$
- Pareto-dominated by both of the pure-strategy equilibria
  - In each of them, one agent gets 1 and the other gets 2

	Husband	Opera	Football
Wife			
Opera		2, 1	0, 0
Football		0, 0	1, 2

# Finding Nash Equilibria

## Matching Pennies

- Each agent has a penny
- Each agent independently chooses to display his/her penny heads up or tails up
- Easy to see that in this game, no pure strategy could be part of a Nash equilibrium
  - For each combination of pure strategies, one of the agents can do better by changing his/her strategy
    - for (Heads,Heads), agent 2 can do better by switching to Tails
    - for (Heads,Tails), agent 1 can do better by switching to Tails
    - for (Tails,Tails), agent 2 can do better by switching to Heads
    - for (Tails,Heads), agent 1 can do better by switching to Heads
- But there's a mixed-strategy equilibrium:
  - $(s,s)$ , where  $s(\text{Heads}) = s(\text{Tails}) = \frac{1}{2}$

Agent 1 \ Agent 2	Heads	Tails
Heads	1, -1	-1, 1
Tails	-1, 1	1, -1

# A Real-World Example

- **Penalty kicks in soccer**

- A kicker and a goalie in a penalty kick
- Kicker can kick left or right
- Goalie can jump to left or right
- Kicker scores iff he/she kicks to one side and goalie jumps to the other
- Analogy to Matching Pennies

- If you use a pure strategy and the other agent uses his/her best response, the other agent will win
- If you kick or jump in either direction with equal probability, the opponent can't exploit your strategy



# Another Interpretation of Mixed Strategies

- Another interpretation of mixed strategies is that
  - Each agent's strategy is deterministic
  - But each agent has uncertainty regarding the other's strategy
- Agent  $i$ 's mixed strategy is **everyone else's assessment** of how likely  $i$  is to play each pure strategy
- Example:
  - In a series of soccer penalty kicks, the kicker could kick left or right in a deterministic pattern that the goalie thinks is random

# Two-Finger Morra

- There are several versions of this game
  - Here's the one the book uses:
- Each agent holds up 1 or 2 fingers
  - If the total number of fingers is odd
    - Agent 1 gets that many points
  - If the total number of fingers is even
    - Agent 2 gets that many points
- Agent 1 has no dominant strategy
  - Agent 2 plays 1  $\Rightarrow$  agent 1's best response is 2
  - Agent 2 plays 2  $\Rightarrow$  agent 1's best response is 1
- Similarly, agent 2 has no dominant strategy
- Thus there's no pure-strategy Nash equilibrium
  - Look for a mixed-strategy equilibrium

Agent 1 \ Agent 2	1 finger	2 fingers
1 finger	-2, 2	3, -3
2 fingers	3, -3	-4, 4



# Two-Finger Morra

Agent 2 Agent 1	1 finger	2 fingers
1 finger	-2, 2	3, -3
2 fingers	3, -3	-4, 4

- Let  $p_1 = P(\text{agent 1 plays 1 finger})$   
and  $p_2 = P(\text{agent 2 plays 1 finger})$
- Suppose  $0 < p_1 < 1$  and  $0 < p_2 < 1$
- If this is a mixed-strategy equilibrium, then
  - 1 finger and 2 fingers must have the same expected utility for agent 1
    - Agent 1 plays 1 finger  $\Rightarrow$  expected utility is  $-2p_2 + 3(1-p_2) = 3 - 5p_2$
    - Agent 1 plays 2 fingers  $\Rightarrow$  expected utility is  $3p_2 - 4(1-p_2) = 7p_2 - 4$
    - Thus  $3 - 5p_2 = 7p_2 - 4$ , so  $p_2 = 7/12$
    - Agent 1's expected utility is  $3 - 5(7/12) = 1/12$
  - 1 finger and 2 fingers must also have the same expected utility for agent 2
    - Agent 2 plays 1 finger  $\Rightarrow$  expected utility is  $2p_1 - 3(1-p_1) = 5p_1 - 3$
    - Agent 2 plays 2 fingers  $\Rightarrow$  expected utility is  $-3p_1 + 4(1-p_1) = 4 - 7p_1$
    - Thus  $5p_1 - 3 = 4 - 7p_1$ , so  $p_1 = 7/12$
    - Agent 2's expected utility is  $5(7/12) - 3 = -1/12$

# Another Real-World Example

- **Road Networks**

- Suppose that 1,000 drivers wish to travel from  $S$  (start) to  $D$  (destination)

- Two possible paths:

- $S \rightarrow A \rightarrow D$  and  $S \rightarrow B \rightarrow D$

- The roads  $S \rightarrow A$  and  $B \rightarrow D$  are very long and very wide

- $t = 50$  minutes for each, no matter how many drivers

- The roads  $S \rightarrow B$  and  $A \rightarrow D$  are very short and very narrow

- Time for each = (number of cars)/25

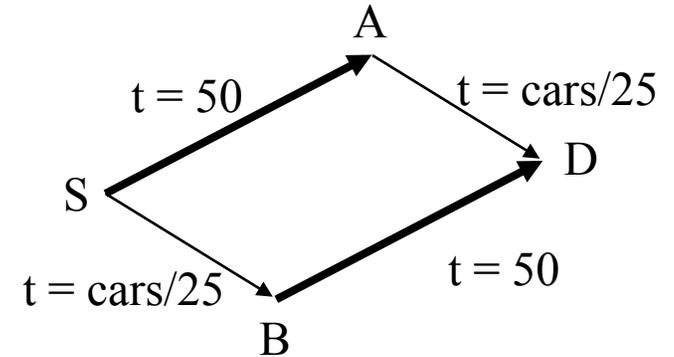
- Nash equilibrium:

- 500 cars go through A, 500 cars through B

- Everyone's time is  $50 + 500/25 = 70$  minutes

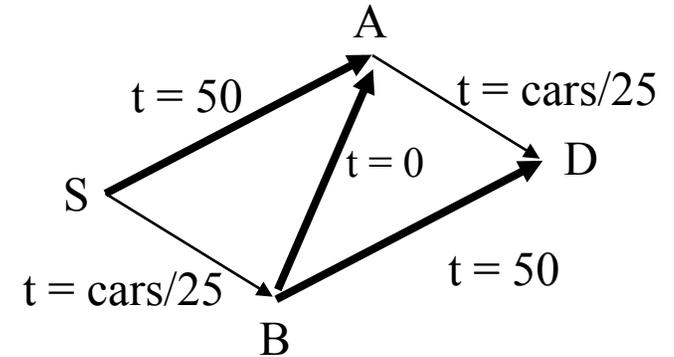
- If a single driver changes to the other route

- There now are 501 cars on that route, so his/her time goes up



# Braess's Paradox

- Suppose we add a new road from B to A
- It's so wide and so short that it takes 0 minutes
- New Nash equilibrium:
  - All 1000 cars go  $S \rightarrow B \rightarrow A \rightarrow D$
  - Time is  $1000/25 + 1000/25 = 80$  minutes
- To see that this is an equilibrium:
  - If driver goes  $S \rightarrow A \rightarrow D$ , his/her cost is  $50 + 40 = 90$  minutes
  - If driver goes  $S \rightarrow B \rightarrow D$ , his/her cost is  $40 + 50 = 90$  minutes
  - Both are dominated by  $S \rightarrow B \rightarrow A \rightarrow D$
- To see that it's the *only* Nash equilibrium:
  - For every traffic pattern, compute the times a driver would get on all three routes
  - In every case,  $S \rightarrow B \rightarrow A \rightarrow D$  dominates  $S \rightarrow A \rightarrow D$  and  $S \rightarrow B \rightarrow D$
- Carelessly adding capacity can actually be hurtful!



# Braess's Paradox in practice

- From an article about Seoul, South Korea:
  - “The idea was sown in 1999,” Hwang says. “We had experienced a strange thing. We had three tunnels in the city and one needed to be shut down. Bizarrely, we found that that car volumes dropped. I thought this was odd. We discovered it was a case of ‘Braess paradox’, which says that by taking away space in an urban area you can actually increase the flow of traffic, and, by implication, by adding extra capacity to a road network you can reduce overall performance.”
- John Vidal, “Heart and soul of the city”, *The Guardian*, Nov. 1, 2006  
<http://www.guardian.co.uk/environment/2006/nov/01/society.travelsenvironmentalimpact>

# The $p$ -Beauty Contest

- Consider the following game:
  - Each player chooses a number in the range from 0 to 100
  - The winner(s) are whoever chose a number that's closest to  $2/3$  of the average
- This game is famous among economists and game theorists
  - It's called the  $p$ -beauty contest
  - I used  $p = 2/3$

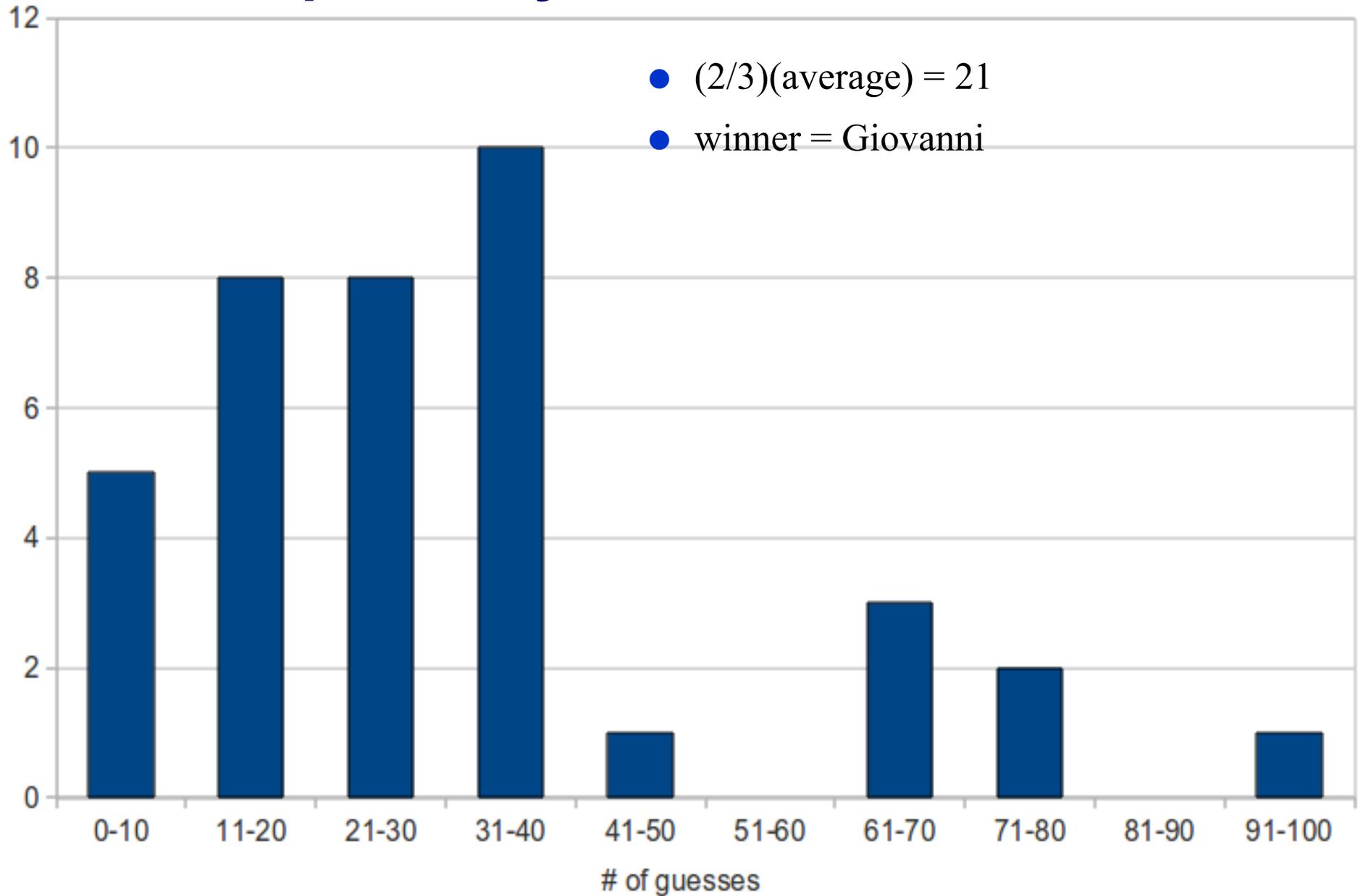
# Elimination of Dominated Strategies

- A strategy  $s_i$  is (**strongly, weakly**) **dominated** for an agent  $i$  if some other strategy  $s'_i$  strictly dominates  $s_i$
- A strictly dominated strategy can't be a best response to any move
  - So we can eliminate it (remove it from the payoff matrix)
- Once a pure strategy is eliminated, another strategy may become dominated
  - This elimination can be repeated

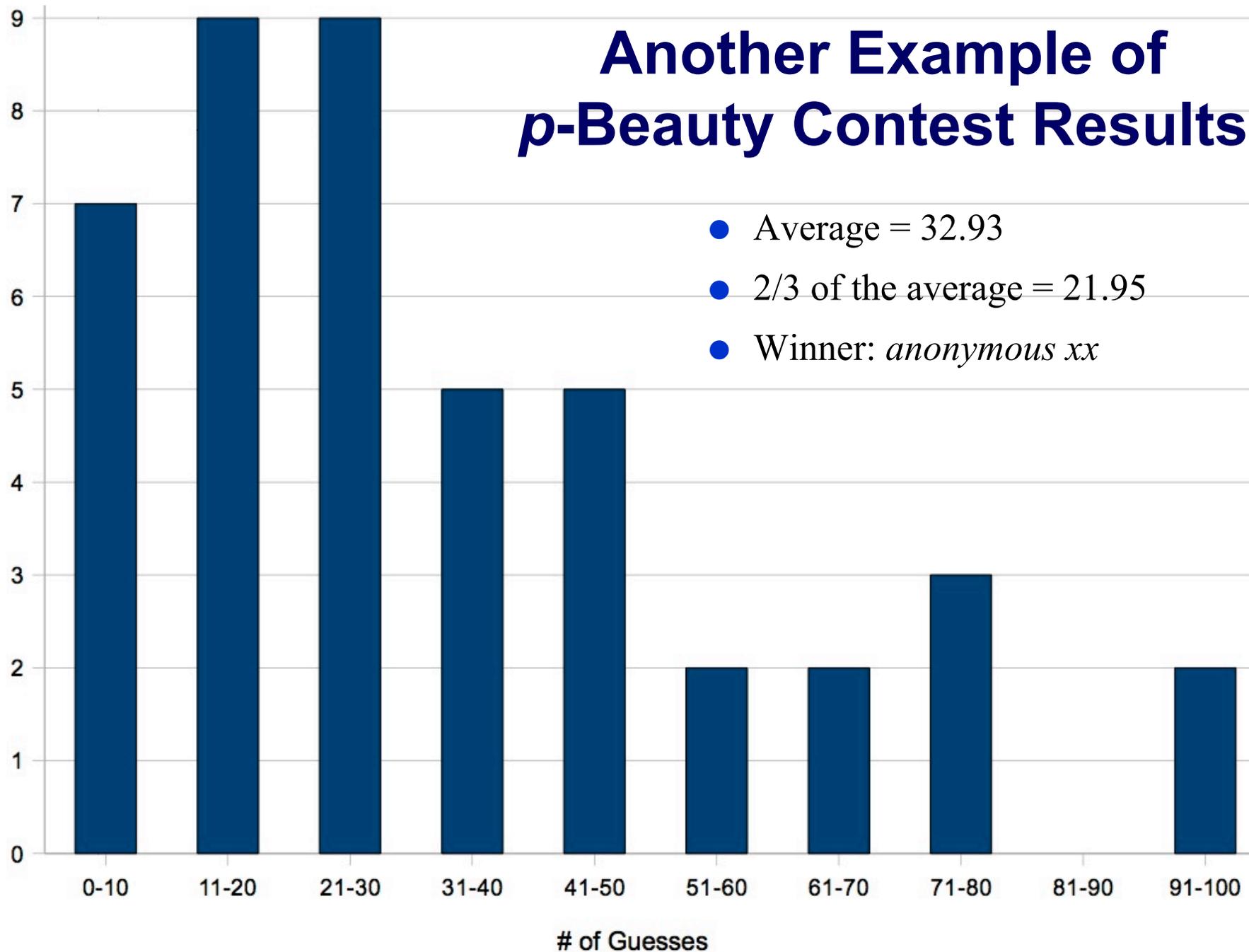
# Iterated Elimination of Dominated Strategies

- Iteratively eliminate strategies that can never be a best response if the other agents play rationally
  - All numbers  $\leq 100 \Rightarrow \frac{2}{3}(\text{average}) < 67$ 
    - $\Rightarrow$  Any rational agent will choose a number  $< 67$
  - All rational choices  $\leq 67 \Rightarrow \frac{2}{3}(\text{average}) < 45$ 
    - $\Rightarrow$  Any rational agent will choose a number  $< 45$
  - All rational choices  $\leq 45 \Rightarrow \frac{2}{3}(\text{average}) < 30$ 
    - $\Rightarrow$  Any rational agent will choose a number  $< 30$
  - ...
- Nash equilibrium: everyone chooses 0

# *p*-Beauty Contest Results



# Another Example of $p$ -Beauty Contest Results



# We aren't rational

- We aren't game-theoretically rational agents
- Huge literature on *behavioral economics* going back to about 1979
  - Many cases where humans (or aggregations of humans) tend to make different decisions than the game-theoretically optimal ones
  - Daniel Kahneman received the 2002 Nobel Prize in Economics for his work on that topic

# Choosing “Irrational” Strategies

- Why choose a non-equilibrium strategy?
  - Limitations in reasoning ability
    - Didn't calculate the Nash equilibrium correctly
    - Don't know how to calculate it
    - Don't even know the concept
  - Hidden payoffs
    - Other things may be more important than winning
      - › Want to be helpful
      - › Want to see what happens
      - › Want to create mischief
  - Agent modeling (next slide)

# Agent Modeling

- A Nash equilibrium strategy is best for you *if the other agents also use their Nash equilibrium strategies*
- In many cases, the other agents *won't* use Nash equilibrium strategies
  - If you can forecast their actions accurately, you may be able to do much better than the Nash equilibrium strategy
- Example: **repeated games**

# Repeated Games

- Used by game theorists, economists, social and behavioral scientists as highly simplified models of various real-world situations



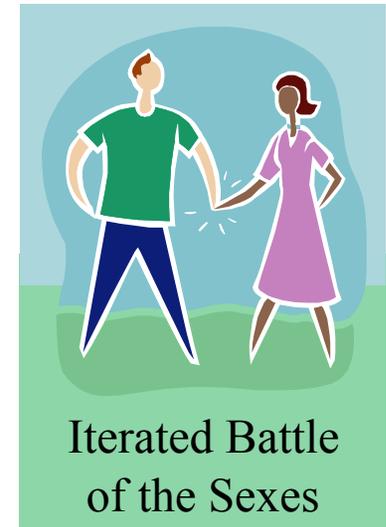
Iterated Prisoner's Dilemma



Roshambo



Repeated  
Ultimatum Game



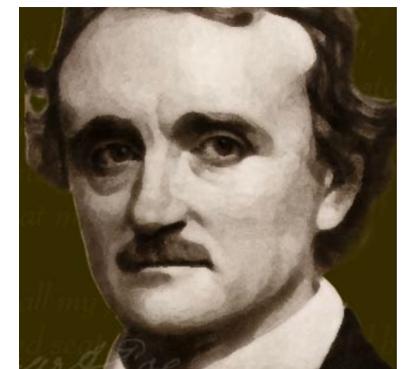
Iterated Battle  
of the Sexes



Iterated Chicken Game



Repeated Stag Hunt



Repeated  
Matching Pennies

# Repeated Games

- In repeated games, some game  $G$  is played multiple times by the same set of agents
  - $G$  is called the **stage game**
  - Each occurrence of  $G$  is called an **iteration** or a **round**
- Usually each agent knows what all the agents did in the previous iterations, but not what they're doing in the current iteration
- Usually each agent's payoff function is additive

Prisoner's Dilemma:

	Agent 2	
Agent 1 \	C	D
C	3, 3	0, 5
D	5, 0	1, 1

Iterated Prisoner's Dilemma, with 2 iterations:

	<i>Agent 1:</i>	<i>Agent 2:</i>
Round 1:	C	C
Round 2:	<i>D</i>	C
Total payoff:	$3+5 = 8$	$3+0 = 3$

# Roshambo (Rock, Paper, Scissors)

$A_2 \backslash A_1$	Rock	Paper	Scissors
Rock	0, 0	-1, 1	1, -1
Paper	1, -1	0, 0	-1, 1
Scissors	-1, 1	1, -1	0, 0



- Nash equilibrium for the stage game:
  - choose randomly,  $P=1/3$  for each move
- Nash equilibrium for the repeated game:
  - *always* choose randomly,  $P=1/3$  for each move
- Expected payoff = 0
- Let's see how that works out in practice ...

# Roshambo (Rock, Paper, Scissors)

$A_2 \backslash A_1$	Rock	Paper	Scissors
Rock	0, 0	-1, 1	1, -1
Paper	1, -1	0, 0	-1, 1
Scissors	-1, 1	1, -1	0, 0

- 1999 international roshambo programming competition  
[www.cs.ualberta.ca/~darse/rsbpc1.html](http://www.cs.ualberta.ca/~darse/rsbpc1.html)
  - Round-robin tournament:
    - 55 programs, 1000 iterations for each pair of programs
    - Lowest possible score = -55000, highest possible score = 55000
  - Average over 25 tournaments:
    - Highest score (*Jocaine Powder*): 13038
    - Lowest score (*Cheesebot*): -36006
  - Very different from the game-theoretic prediction



# Opponent Modeling

- A Nash equilibrium strategy is best for you  
*if the other agents also use their Nash equilibrium strategies*
- In many cases, the other agents *won't* use Nash equilibrium strategies
  - If you can forecast their actions accurately, you may be able to do much better than the Nash equilibrium strategy
- One reason why the other agents might not use their Nash equilibrium strategies:
  - Because they may be trying to forecast *your* actions too

# Iterated Prisoner's Dilemma

- Multiple iterations of the *Prisoner's Dilemma*.

- Widely used to study the emergence of cooperative behavior among agents
  - e.g., Axelrod (1984), *The Evolution of Cooperation*
- Axelrod ran a famous set of tournaments
  - People contributed strategies encoded as computer programs
  - Axelrod played them against each other

Prisoner's Dilemma:

	Agent 2	C	D
Agent 1			
C		3, 3	0, 5
D		5, 0	1, 1

Nash equilibrium

If I defect now, he might punish me by defecting next time



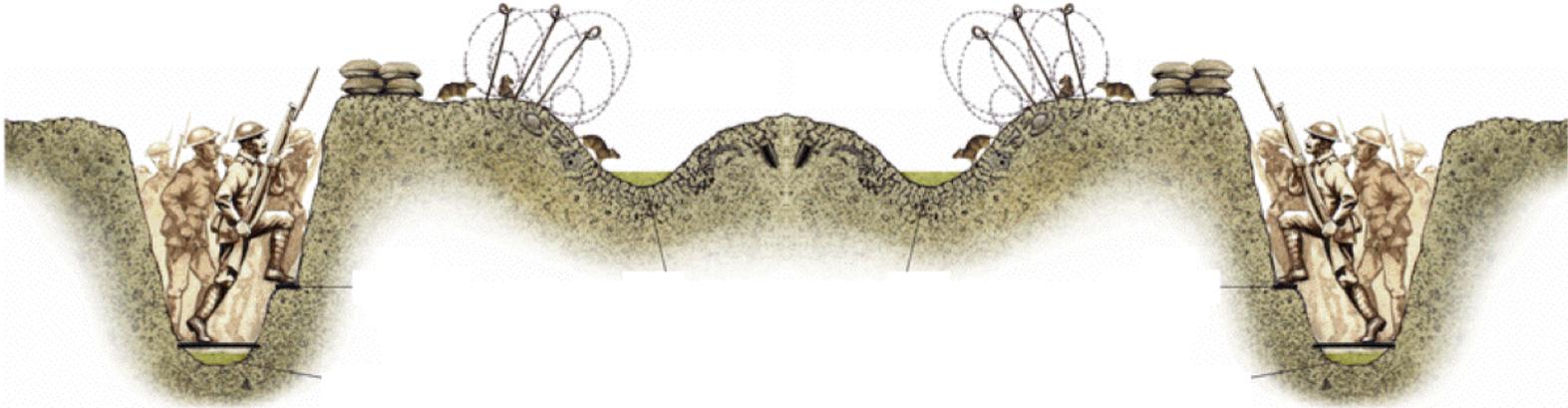
# TFT with Other Agents

- In Axelrod's tournaments, TFT usually did best
  - » It could establish and maintain cooperations with many other agents
  - » It could prevent malicious agents from taking advantage of it

<i>TFT AllC</i>	<i>TFT AllD</i>	<i>TFT Grim</i>	<i>TFT TFT</i>	<i>TFT Tester</i>
C C	C <i>D</i>	C C	C C	C <i>D</i>
C C	<i>D D</i>	C C	C C	<i>D</i> C
C C	<i>D D</i>	C C	C C	C C
C C	<i>D D</i>	C C	C C	C C
C C	<i>D D</i>	C C	C C	C C
C C	<i>D D</i>	C C	C C	C C
C C	<i>D D</i>	C C	C C	C C
⋮ ⋮	⋮ ⋮	⋮ ⋮	⋮ ⋮	⋮ ⋮

# Example:

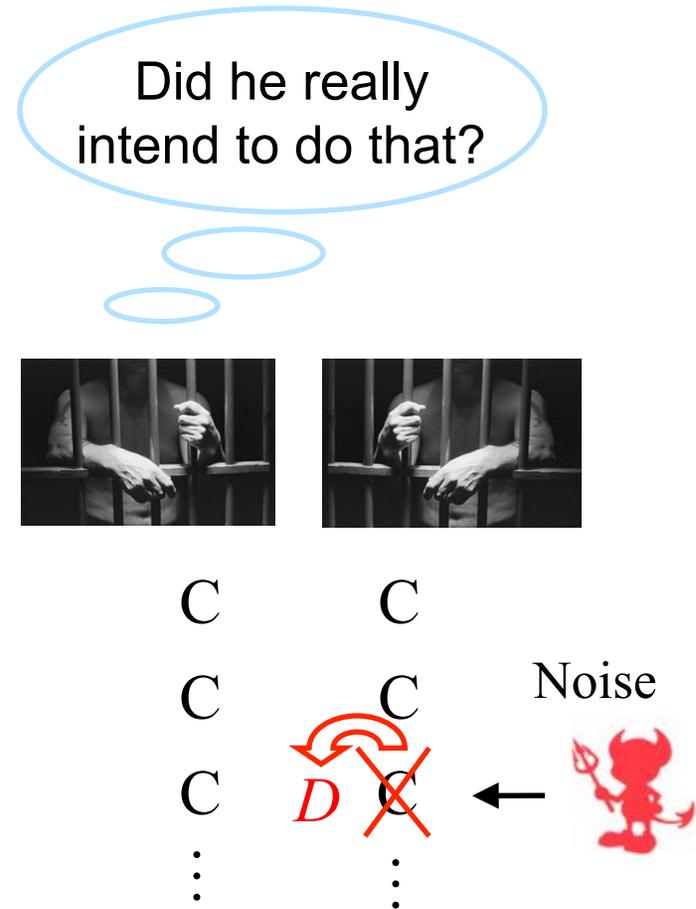
- A real-world example of the IPD, described in Axelrod's book:
  - World War I trench warfare



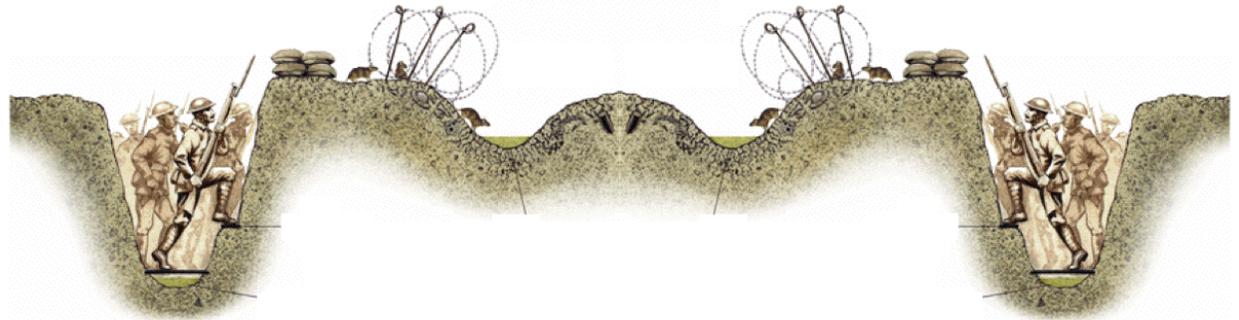
- Incentive to cooperate:
  - If I attack the other side, then they'll retaliate and I'll get hurt
  - If I don't attack, maybe they won't either
- Result: evolution of cooperation
  - Although the two infantries were supposed to be enemies, they avoided attacking each other

# IPD with Noise

- In noisy environments,
  - There's a nonzero probability (e.g., 10%) that a "noise gremlin" will change some of the actions
    - *Cooperate* (C) becomes *Defect* (D), and vice versa
- Can use this to model accidents
  - Compute the score using the changed action
- Can also model misinterpretations
  - Compute the score using the original action



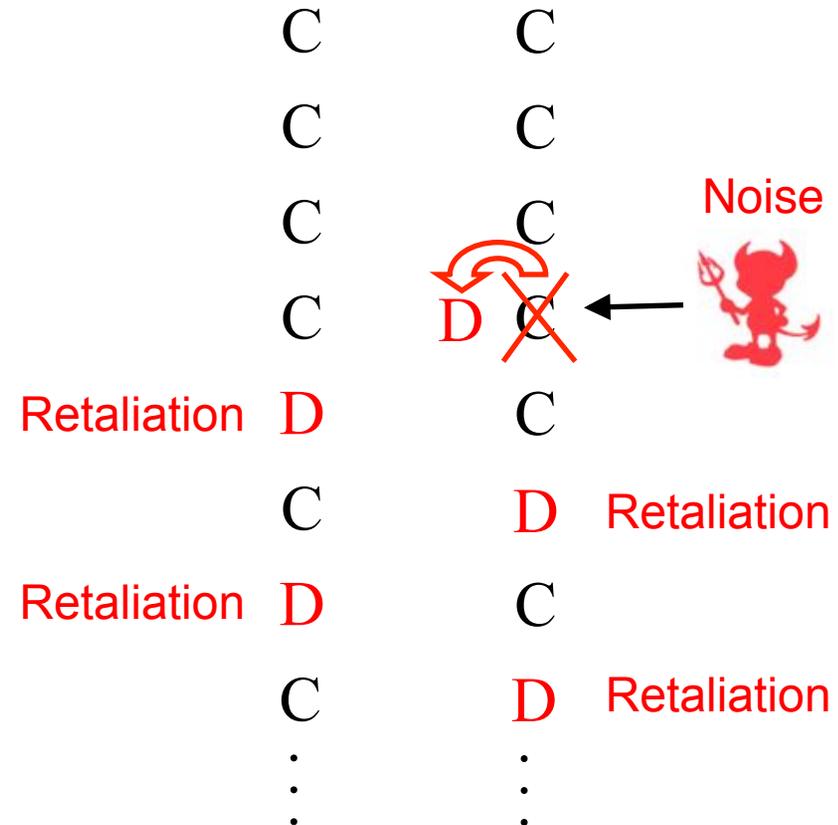
# Example of Noise



- Story from a British army officer in World War I:
  - I was having tea with A Company when we heard a lot of shouting and went out to investigate. We found our men and the Germans standing on their respective parapets. **Suddenly a salvo arrived but did no damage.** Naturally both sides got down and our men started swearing at the Germans, when all at once **a brave German got onto his parapet and shouted out: “We are very sorry about that; we hope no one was hurt. It is not our fault. It is that damned Prussian artillery.”**
- The salvo wasn't the German infantry's intention
  - They didn't expect it nor desire it

# Noise Makes it Difficult to Maintain Cooperation

- Consider two agents who both use TFT
- One accident or misinterpretation can cause a long string of retaliations



# Some Strategies for the Noisy IPD

- **Principle:** be more forgiving in the face of defections
- Tit-For-Two-Tats (TFTT)
  - › Retaliate only if the other agent defects twice in a row
    - Can tolerate isolated instances of defections, but susceptible to exploitation of its generosity
    - Beaten by the TESTER strategy I described earlier
- Generous Tit-For-Tat (GTFT)
  - › Forgive randomly: small probability of cooperation if the other agent defects
  - › Better than TFTT at avoiding exploitation, but worse at maintaining cooperation
- Pavlov
  - › Win-Stay, Lose-Shift
    - Repeat previous move if I earn 3 or 5 points in the previous iteration
    - Reverse previous move if I earn 0 or 1 points in the previous iteration
  - › Thus if the other agent defects continuously, Pavlov will alternatively cooperate and defect

# Discussion

- The British army officer's story:
  - a German shouted, ``We are very sorry about that; we hope no one was hurt. It is not our fault. It is that damned Prussian artillery.''
- The apology avoided a conflict
  - It was convincing because it was consistent with the German infantry's past behavior
  - The British had ample evidence that the German infantry wanted to keep the peace
- If you can tell which actions are *affected* by noise, you can avoid *reacting* to the noise
- IPD agents often behave deterministically
  - For others to cooperate with you it helps if you're predictable
- This makes it feasible to build a model from observed behavior

# The DBS Agent

- Work by my recent PhD graduate, Tsz-Chiu Au
  - Now a postdoc at University of Texas
- From the other agent's recent behavior, build a model  $\pi$  of the other agent's strategy
  - A set of rules giving the probability of each action in various situations
- Use the model to filter noise
  - Observed move contradicts the model  $\Rightarrow$  assume the observed move is noise
- Detect changes in the other agent's strategy
  - Observed move contradicts the model too many times  $\Rightarrow$  assume they've changed their strategy; recompute the model
- Use the model to help plan our next move
  - Game-tree search, using  $\pi$  to predict the other agent's moves

Au & Nau. Accident or intention: That is the question (in the iterated prisoner's dilemma). *AAMAS*, 2006.

Au & Nau. Is it accidental or intentional? A symbolic approach to the noisy iterated prisoner's dilemma. In G. Kendall (ed.), *The Iterated Prisoners Dilemma: 20 Years On*. World Scientific, 2007.

# 20<sup>th</sup> Anniversary IPD Competition

<http://www.prisoners-dilemma.com>

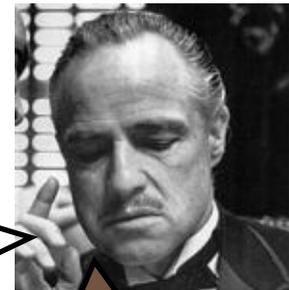
- Category 2: IPD with noise
  - 165 programs participated
- DBS dominated the top 10 places
- Two agents scored higher than DBS
  - They both used *master-and-slaves* strategies

Rank	Program	Avg. score
1	BWIN	433.8
2	IMM01	414.1
3	DBSz	408.0
4	DBSy	408.0
5	DBSpl	407.5
6	DBSx	406.6
7	DBSf	402.0
8	DBStft	401.8
9	DBSd	400.9
10	lowESTFT_classic	397.2
11	TFTIm	397.0
12	Mod	396.9
13	TFTIz	395.5
14	TFTIc	393.7
15	DBSe	393.7
16	TTFT	393.4
17	TFTIa	393.3
18	TFTIb	393.1
19	TFTIx	393.0
20	mediumESTFT_classic	392.9

# Master & Slaves Strategy

- Each participant could submit up to 20 programs
- Some submitted programs that could recognize each other
  - (by communicating pre-arranged sequences of Cs and Ds)
- The 20 programs worked as a team
  - 1 master, 19 slaves
  - When a slave plays with its master
    - Slave cooperates, master defects
    - => maximizes the master's payoff
  - When a slave plays with an agent not in its team
    - It defects
    - => minimizes the other agent's payoff

My goons give me all their money ...



... and they beat up everyone else



# Comparison

- Analysis
  - Each master-slaves team's average score was much lower than DBS's
  - If BWIN and IMM01 had each been restricted to  $\leq 10$  slaves, DBS would have placed 1st
  - Without any slaves, BWIN and IMM01 would have done badly
- In contrast, DBS had no slaves
  - DBS established cooperation with *many* other agents
  - DBS did this *despite* the noise, because it filtered out the noise



# Summary

- Dominant strategies and dominant strategy equilibria
  - Prisoner's dilemma
- Pareto optimality
- Best responses and Nash equilibria
  - Battle of the Sexes, Matching Pennies, Two-Finger Morra
- Real-world examples
  - Soccer penalty kicks, road networks (Braess's Paradox)
- Repeated games and opponent modeling
  - roshambo (rock-paper-scissors)
  - iterated prisoner's dilemma with noise
    - opponent models based on observed behavior
    - detection and removal of noise, game-tree search
  - 20<sup>th</sup> anniversary IPD competition