# 07-380 Notes
# Pre-reading: Game Theory

Carnegie Mellon University
School of Computer Science

## Contents

**1 Summary** **2**

**2 Game Theory** **2**

**3 Games** **3**
3.1 Payoff Matrices . . . 3
3.2 Normal-Form Games . . . 3
3.3 Two More Games . . . 4
3.4 Zero-Sum and General-Sum Games . . . 5

**4 Strategies** **5**
4.1 Pure and Mixed Strategies . . . 5
4.2 Profiles: Plain for One Player, Bold for All Players . . . 6

**5 Expected Utility** **6**
5.1 The Probability of an Action Profile . . . 6
5.2 Expected Utility . . . 7

**6 Best Response** **8**
6.1 Everyone Except Player $i$ . . . 8
6.2 Definition . . . 8
6.3 Example: Best Response to an Action . . . 9
6.4 Example: Best Response to a Mixed Strategy . . . 9

**A Additional Examples** **10**
A.1 Questions . . . 10
A.2 Solutions . . . 11

**B Game Theory: People and Applications** **13**
B.1 John Nash . . . 13
B.2 Auctions . . . 13
B.3 Security and Wildlife Protection . . . 14
B.4 Poker . . . 14
B.5 Machine Learning . . . 15
B.6 Sources . . . 15

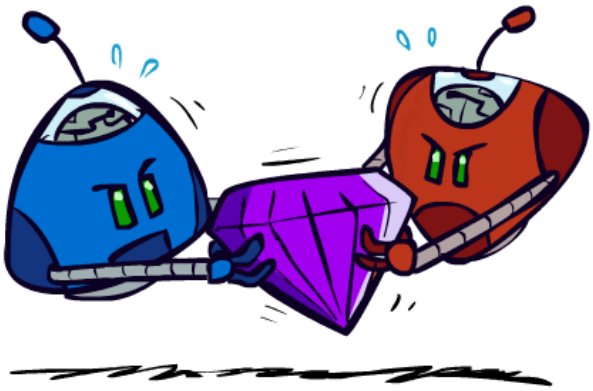


Illustration: Ketrina Yim / UC Berkeley AI

# 1 Summary

This pre-reading introduces the vocabulary and the notation of game theory that our lecture builds on. Everything is collected here for reference. Sections 3 to 6 explain it step by step.

**The three games** that we use in these notes and in lecture are below. Everything that belongs to the row player is blue, and everything that belongs to the column player is green.

| Player 1 \ Player 2 | 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**

| Alex \ Barrie | Football | Concert |
|---|---|---|
| Football | 2, 1 | 0, 0 |
| Concert | 0, 0 | 1, 2 |

**Football vs. Concert**

| Player 1 \ Player 2 | Cooperate | Defect |
|---|---|---|
| Cooperate | −1, −1 | −3, 0 |
| Defect | 0, −3 | −2, −2 |

**Prisoner's Dilemma**

**Notation**:

| Notation | Meaning | Example |
|---|---|---|
| $a_i$ | An action of player $i$ | Rock |
| $\mathbf{a} = (a_1, a_2)$ | Action profile: one action per player. It is one cell of the payoff matrix | (Paper, Rock) |
| $u_i(\mathbf{a})$ | Utility (payoff) of player $i$ for an action profile | $u_2(\text{Paper}, \text{Rock}) = -1$ |
| $s_i$ | Strategy of player $i$: one probability per action | $s_1 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$ |
| Pure strategy | A strategy that always plays the same action | $(0, 1, 0)$, or just Paper |
| Mixed strategy | A strategy that randomizes over actions | $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$ |
| Support | The actions that a strategy plays with non-zero probability | {Rock, Paper} |
| $\mathbf{s} = (s_1, s_2)$ | Strategy profile: one strategy per player | $\left(\left(\frac{1}{2}, \frac{1}{2}, 0\right), (0, 1, 0)\right)$ |
| $u_i(\mathbf{s})$ | Expected utility of player $i$ for a strategy profile | $\sum_{\mathbf{a}} P(\mathbf{a})\, u_i(\mathbf{a})$ |
| $\mathbf{s}_{-i}$ | The strategies of everyone except player $i$ | $\mathbf{s}_{-1} = s_2$ |
| $\text{BR}(\mathbf{s}_{-i})$ | Best response: the strategies of player $i$ with the highest expected utility $u_i(s_i, \mathbf{s}_{-i})$ | $\text{Paper} \in \text{BR}(\text{Rock})$ |

# 2 Game Theory

In 07-280, we used adversarial search to choose moves in games like chess, where two players take turns and one player's win is the other player's loss.

Many situations with more than one decision maker do not look like that. Two companies set their prices at the same time. Every driver picks a route, and the traffic depends on what all of them picked. How well a choice works out depends on what everyone else chooses.

**Game theory** is the study of mathematical models of conflict and cooperation between intelligent decision makers.

A **game**, in this sense, is any situation in which the result for each decision maker depends on the actions of all of them.

The idea at the center of our lecture, the Nash equilibrium, is named after John Nash, who studied mathematics at Carnegie Tech before it became Carnegie Mellon. Appendix B has more about Nash and about where game theory is used today, including work by CMU researchers on poker and on protecting wildlife.

# 3 Games

## 3.1 Payoff Matrices

Our first game is Rock-Paper-Scissors. Two players each choose Rock, Paper, or Scissors at the same time. Rock beats Scissors, Scissors beats Paper, and Paper beats Rock.

Each player gets a **utility**, a number that says how much the player likes the outcome. The winner gets 1, the loser gets $-1$, and in a tie both get 0. A utility is also called a **payoff**.

The whole game fits in one table, called a **payoff matrix**:

| Player 1 \ Player 2 | 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 1 is the **row player**. Each row is one of Player 1's actions.
- Player 2 is the **column player**. Each column is one of Player 2's actions.
- Each cell holds two numbers. The first is Player 1's utility and the second is Player 2's utility when that row and that column are played.

For example, suppose Player 1 plays Paper and Player 2 plays Rock. In the Paper row and the Rock column (the shaded cell) we find $1, -1$. Paper beats Rock, so Player 1 gets 1 and Player 2 gets $-1$.

## 3.2 Normal-Form Games

Rock-Paper-Scissors is a game in **normal form**. A normal-form game consists of:

- A set of **players**. We number them $1, 2, \ldots, N$ and write $i$ for any one of them.
- A set of **actions** for each player. We write $a_i$ for an action of player $i$.
- A **utility function** $u_i$ for each player. It gives player $i$'s utility for every combination of the actions chosen by *all* of the players.

The players move **simultaneously**: each player chooses an action without knowing what the others have chosen. Then every player receives their utility, and the game is over.

One action for each player, taken together, is called an **action profile**. We write it in bold:

$$\mathbf{a} = (a_1, a_2, \ldots, a_N).$$

Each cell of a payoff matrix is one action profile. The shaded cell above is $\mathbf{a} = (\text{Paper}, \text{Rock})$, and its two numbers are

$$u_1(\text{Paper}, \text{Rock}) = 1 \qquad \text{and} \qquad u_2(\text{Paper}, \text{Rock}) = -1.$$

Two conventions to hold on to:

- The order inside the parentheses is always (Player 1's action, Player 2's action), no matter whose utility we are asking for.
- The subscript on $u$ says whose utility it is, and so which of the two numbers in the cell to read.

A normal-form game with two players and a finite number of actions is called a **bimatrix game**. These are the games that fit in a payoff matrix.

## 3.3 Two More Games

**Football vs. Concert**: Alex and Barrie want to spend the evening together, but each has to decide alone whether to go to the football game or to the concert. Alex prefers football and Barrie prefers the concert, but both would rather be together than alone.

- Together at the football game: Alex gets 2 and Barrie gets 1.
- Together at the concert: Alex gets 1 and Barrie gets 2.
- At different events: both get 0.

| | | Barrie: Football | Barrie: Concert |
|---|---|---|---|
| Alex | Football | 2, 1 | 0, 0 |
| | Concert | 0, 0 | 1, 2 |

With named players we use their initials as subscripts. For example, $u_A(\text{Football}, \text{Concert}) = 0$ and $u_B(\text{Concert}, \text{Concert}) = 2$.

**Prisoner's Dilemma**: Two members of a gang are arrested and held in separate rooms. Each must choose to *Cooperate* with their partner by staying silent, or to *Defect* by testifying against their partner.

- Both Cooperate: each gets 1 year in jail.
- Both Defect: each gets 2 years.
- Only one Defects: that prisoner goes free, and the other gets 3 years.

A player's utility is the negative of their years in jail.

| | | Player 2: Cooperate | Player 2: Defect |
|---|---|---|---|
| Player 1 | Cooperate | −1, −1 | −3, 0 |
| | Defect | 0, −3 | −2, −2 |

### 3.4 Zero-Sum and General-Sum Games

In a **zero-sum game**, the utilities of all the players add up to zero no matter which actions are chosen. (Adding up to the same constant for every action profile also counts.)

In a **general-sum game**, the sum of the utilities is not constant. It depends on which actions are chosen.

To tell which kind a game is, add the two numbers in each cell:

- **Rock-Paper-Scissors**: every cell adds up to 0. Zero-sum.
- **Football vs. Concert**: the cells add up to 3, 0, 0, and 3. General-sum.
- **Prisoner's Dilemma**: the cells add up to $-2$, $-3$, $-3$, and $-4$. General-sum.

A zero-sum game is pure conflict: whatever one player gains, the other loses. In a general-sum game, the players' interests can be partly aligned. Alex and Barrie both prefer being together to being apart.

## 4 Strategies

### 4.1 Pure and Mixed Strategies

A player's **strategy** says how the player chooses an action.

- A **pure strategy** chooses one action deterministically: "play Rock."
- A **mixed strategy** chooses an action at random, according to a probability distribution: "play Rock with probability 0.3, Paper with probability 0.7, and never play Scissors."

We write a strategy as a list of probabilities, one for each action, in the order that the actions appear in the payoff matrix. With the order (Rock, Paper, Scissors), the mixed strategy above is

$$s = (0.3,\ 0.7,\ 0).$$

The probabilities add up to one.

The **support** of a strategy is the set of actions that it plays with non-zero probability. The support of $(0.3, 0.7, 0)$ is $\{\text{Rock}, \text{Paper}\}$.

A pure strategy is the special case in which one action has all of the probability. "Play Rock" is the strategy $(1, 0, 0)$, so we will often write an action where a strategy is expected.

**Notation alert**: In search and in MDPs, $s$ was a state. In game theory, $s$ is a strategy.

**Why randomize?** If we always play Rock, an opponent who figures that out will always play Paper. Being unpredictable can be the best thing that we can do.

### 4.2 Profiles: Plain for One Player, Bold for All Players

Every player has an action, a strategy, and a utility. The notation follows one rule: a plain letter with a subscript belongs to one player, and a bold letter is a **profile**, the list of that thing for all of the players.

| | One player | All players (a profile) |
|---|---|---|
| Action | $a_i$ | $\mathbf{a} = (a_1, a_2, \ldots, a_N)$ |
| Strategy | $s_i$ | $\mathbf{s} = (s_1, s_2, \ldots, s_N)$ |
| Utility | $u_i$ | $\mathbf{u} = (u_1, u_2, \ldots, u_N)$ |

For example, if Player 1 uses $s_1 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$ and Player 2 always plays Paper, $s_2 = (0, 1, 0)$, then the **strategy profile** is the pair of them:

$$\mathbf{s} = (s_1, s_2) = \Big(\left(\tfrac{1}{2}, \tfrac{1}{2}, 0\right),\ (0, 1, 0)\Big).$$

## 5 Expected Utility

### 5.1 The Probability of an Action Profile

When the players use mixed strategies, the action profile that gets played is random. Each player randomizes on their own, so the two choices are independent and their probabilities multiply:

$$P(\mathbf{a}) = P(a_1) \cdot P(a_2).$$

Here $P(a_1)$ is the probability that Player 1's strategy puts on action $a_1$, and $P(a_2)$ is the probability that Player 2's strategy puts on action $a_2$.

**Example**: In Football vs. Concert, suppose Alex flips a fair coin and Barrie always goes to the football game. With the actions in the order (Football, Concert), their strategies are

$$s_A = \left(\tfrac{1}{2}, \tfrac{1}{2}\right) \qquad \text{and} \qquad s_B = (1, 0).$$

We write each player's probabilities next to their actions: Alex's beside the rows and Barrie's above the columns. The probability of a cell is its row probability times its column probability. It is shown on the second line of each cell.

| | | Barrie | |
|---|---|---|---|
| | | 1 | 0 |
| | | Football | Concert |
| Alex $\frac{1}{2}$ | Football | 2, 1<br>$\frac{1}{2} \cdot 1 = \frac{1}{2}$ | 0, 0<br>$\frac{1}{2} \cdot 0 = 0$ |
| Alex $\frac{1}{2}$ | Concert | 0, 0<br>$\frac{1}{2} \cdot 1 = \frac{1}{2}$ | 1, 2<br>$\frac{1}{2} \cdot 0 = 0$ |

Barrie never goes to the concert, so both cells in the Concert column have probability zero (shaded gray). The four probabilities add up to one.

## 5.2 Expected Utility

Since the action profile is random, so is each player's utility. What a player cares about is their utility on average.

Recall the expected value from the 07-280 notes. It is a weighted average of the values, weighted by their probabilities:

$$\mathbb{E}_X[f(X)] = \sum_x P(X = x)\, f(x).$$

Here the random outcome is the action profile $\mathbf{a}$, and the value that we average is player $i$'s utility for it. The **expected utility** of player $i$ for the strategy profile $\mathbf{s}$ is

$$u_i(\mathbf{s}) = \sum_{\mathbf{a}} P(\mathbf{a})\; u_i(\mathbf{a}).$$

The sum has one term for every cell of the payoff matrix: the probability of the cell times player $i$'s utility in the cell.

**Example, continued**: We already have the probability of each cell. For Alex's expected utility, we multiply each one by Alex's utility in that cell, the first number, and add. Going through the four cells in reading order, with F for Football and C for Concert:

$$u_A(s_A, s_B) = \underbrace{\tfrac{1}{2}}_{P(\mathrm{F},\mathrm{F})} \cdot 1 \cdot 2 \;+\; \underbrace{\tfrac{1}{2}}_{P(\mathrm{F},\mathrm{C})} \cdot 0 \cdot 0 \;+\; \underbrace{\tfrac{1}{2}}_{P(\mathrm{C},\mathrm{F})} \cdot 1 \cdot 0 \;+\; \underbrace{\tfrac{1}{2}}_{P(\mathrm{C},\mathrm{C})} \cdot 0 \cdot 1 \;=\; 1.$$

For Barrie's expected utility, the probabilities are exactly the same. Only the last number in each term changes, to Barrie's utility, the second number in the cell:

$$u_B(s_A, s_B) = \tfrac{1}{2} \cdot 1 \cdot 1 \;+\; \tfrac{1}{2} \cdot 0 \cdot 0 \;+\; \tfrac{1}{2} \cdot 1 \cdot 0 \;+\; \tfrac{1}{2} \cdot 0 \cdot 2 \;=\; \tfrac{1}{2}.$$

This matches the story. Barrie is always at the football game. Half of the time Alex is there too, which is worth 2 to Alex and 1 to Barrie. The other half of the time Alex is alone at the concert, which is worth 0 to both.

**Skip the zeros**: A term is zero whenever its cell has probability zero or its utility is zero. In both sums above, only the first term survived.

**One symbol, two uses**: Given an action profile, $u_i(\mathbf{a})$ is a number that we read from the payoff matrix. Given a strategy profile, $u_i(\mathbf{s})$ is an expected utility that we compute with the sum. We can also mix the two, as in $u_A(\text{Football}, s_B)$: Alex's expected utility when Alex plays Football and Barrie uses the strategy $s_B$.

# 6 Best Response

So far, we were given a strategy for every player and computed utilities. Now we start to *choose* strategies.

The first step is a simplified question. Suppose that we know the strategies of all the other players. What should we do?

## 6.1 Everyone Except Player $i$

We need a name for "the strategies of all the other players." Starting from a strategy profile $\mathbf{s}$, we write $\mathbf{s}_{-i}$ for the same list with player $i$'s strategy left out. Read the subscript $-i$ as "everyone except $i$." With three players,

$$\mathbf{s} = (s_1, s_2, s_3), \qquad \mathbf{s}_{-1} = (s_2, s_3), \qquad \mathbf{s}_{-2} = (s_1, s_3).$$

With two players, it is simply the other player's strategy: $\mathbf{s}_{-1} = s_2$ and $\mathbf{s}_{-2} = s_1$.

To put the spotlight on player $i$, we write player $i$'s expected utility as $u_i(s_i, \mathbf{s}_{-i})$: what player $i$ does, and what everyone else does.

## 6.2 Definition

Given the strategies of all the other players, $\mathbf{s}_{-i}$, player $i$'s **best response** to $\mathbf{s}_{-i}$ is the set of strategies for player $i$ that give player $i$ the highest expected utility:

$$\mathrm{BR}(\mathbf{s}_{-i}) = \underset{s_i}{\operatorname{argmax}}\ u_i(s_i, \mathbf{s}_{-i}).$$

Three things to notice:

- **The other players are held fixed.** We are not yet asking what they *should* do.
- **Only player $i$'s utility matters.** The other players' utilities play no part.
- **It is a set.** If several strategies tie for the highest expected utility, all of them are best responses. Question 3 in Appendix A is an example.

In a payoff matrix, which numbers we compare depends on whose best response we want:

- For the row player, the column player's strategy is fixed. The row player picks a row, so we compare the first numbers from one *row* to another.
- For the column player, the row player's strategy is fixed. The column player picks a column, so we compare the second numbers from one *column* to another.

## 6.3 Example: Best Response to an Action

In Rock-Paper-Scissors, Player 2 plays Rock. What is Player 1's best response?

Player 2's action is fixed, so we stay in the Rock column (shaded). Player 1 chooses the row, so we compare Player 1's utilities, the first numbers, down that column (underlined).

| Player 1 \ Player 2 | Rock | Paper | Scissors |
|---|---|---|---|
| Rock | $\underline{0}, 0$ | $-1, 1$ | $1, -1$ |
| Paper | $\boxed{\underline{1}}, -1$ | $0, 0$ | $-1, 1$ |
| Scissors | $\underline{-1}, 1$ | $1, -1$ | $0, 0$ |

They are $0$ for Rock, $1$ for Paper, and $-1$ for Scissors. The highest is $1$, so Player 1's best response to Rock is Paper.

In strategy notation, Player 2's strategy is $s_2 = (1, 0, 0)$, and Player 1's best response to it is $s_1 = (0, 1, 0)$.

## 6.4 Example: Best Response to a Mixed Strategy

In Football vs. Concert, Barrie flips a fair coin, $s_B = \left(\frac{1}{2}, \frac{1}{2}\right)$. What is Alex's best response?

This time Alex does not know which column will be played, so there is no single column to look down. Instead, we compute Alex's *expected* utility for each of Alex's actions and compare.

Each one is a weighted average of Alex's utilities along one row, weighted by Barrie's probabilities:

| Alex \ Barrie | Football ($\frac{1}{2}$) | Concert ($\frac{1}{2}$) | |
|---|---|---|---|
| Football | $\underline{2}, 1$ | $\underline{0}, 0$ | $\rightarrow u_A(\text{Football}, s_B) = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 0 = \boxed{1}$ |
| Concert | $\underline{0}, 0$ | $\underline{1}, 2$ | $\rightarrow u_A(\text{Concert}, s_B) = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = \frac{1}{2}$ |

Since $1 > \frac{1}{2}$, Alex's best response to $s_B$ is Football.

A mixed strategy cannot do better for Alex. If Alex plays Football with probability $p$ and Concert with probability $1 - p$, Alex's expected utility is $p \cdot 1 + (1 - p) \cdot \frac{1}{2}$, an average of the two numbers above. That average is highest when $p = 1$.

**To find a best response**: compute one expected utility for each of the player's actions, with the other players' strategies fixed, and pick the action with the highest one.

In lecture, we use best responses to ask what all of the players should do at once. The answer is a strategy profile in which every player's strategy is a best response to everyone else's: a Nash equilibrium.

# A Additional Examples

Three more examples for practice. The questions are first, and each one comes with everything needed to answer it. The solutions start on the next page.

## A.1 Questions

### Question 1: Expected utility with two mixed strategies

Two players play Rock-Paper-Scissors. Player 1 chooses the row and Player 2 chooses the column. Each cell lists Player 1's utility first and Player 2's utility second.

| Player 1 \ Player 2 | Rock | Paper | Scissors |
|---|---|---|---|
| Rock | 0, 0 | −1, 1 | 1, −1 |
| Paper | 1, −1 | 0, 0 | −1, 1 |
| Scissors | −1, 1 | 1, −1 | 0, 0 |

With the actions in the order (Rock, Paper, Scissors), the players use the strategies

$$s_1 = \left(\tfrac{1}{2}, \tfrac{1}{2}, 0\right) \quad \text{and} \quad s_2 = \left(0, \tfrac{1}{3}, \tfrac{2}{3}\right).$$

(a) The sum for an expected utility has nine terms, one for each cell. How many of them are non-zero?

(b) What is $u_1(s_1, s_2)$?

(c) What is $u_2(s_1, s_2)$?

### Question 2: Best responses in the Prisoner's Dilemma

Two prisoners each choose to Cooperate or to Defect. A player's utility is the negative of that player's years in jail. Player 1 chooses the row and Player 2 chooses the column. Each cell lists Player 1's utility first and Player 2's utility second.

| Player 1 \ Player 2 | Cooperate | Defect |
|---|---|---|
| Cooperate | −1, −1 | −3, 0 |
| Defect | 0, −3 | −2, −2 |

(a) What is Player 1's best response if Player 2 Cooperates?

(b) What is Player 1's best response if Player 2 Defects?

### Question 3: When best responses tie

Two players play Rock-Paper-Scissors again. Player 1 chooses the row and Player 2 chooses the column. Each cell lists Player 1's utility first and Player 2's utility second.

| Player 1 \ Player 2 | Rock | Paper | Scissors |
|---|---|---|---|
| Rock | 0, 0 | −1, 1 | 1, −1 |
| Paper | 1, −1 | 0, 0 | −1, 1 |
| Scissors | −1, 1 | 1, −1 | 0, 0 |

With the actions in the order (Rock, Paper, Scissors), Player 1 uses the strategy

$$s_1 = \left(0, \tfrac{1}{3}, \tfrac{2}{3}\right).$$

(a) Which of Player 2's three actions are best responses to $s_1$?

(b) Is the mixed strategy $s_2 = \left(\tfrac{1}{2}, 0, \tfrac{1}{2}\right)$ a best response to $s_1$?

## A.2 Solutions

**Solution 1: Expected utility with two mixed strategies**

Write the probabilities beside the rows and above the columns, and multiply to get the probability of each cell.

| Player 1 \ Player 2 | 0 | $\frac{1}{3}$ | $\frac{2}{3}$ |
|---|---|---|---|
| | Rock | Paper | Scissors |
| $\frac{1}{2}$ Rock | 0, 0<br>0 | −1, 1<br>$\frac{1}{2}\cdot\frac{1}{3}=\frac{1}{6}$ | 1, −1<br>$\frac{1}{2}\cdot\frac{2}{3}=\frac{1}{3}$ |
| $\frac{1}{2}$ Paper | 1, −1<br>0 | 0, 0<br>$\frac{1}{2}\cdot\frac{1}{3}=\frac{1}{6}$ | −1, 1<br>$\frac{1}{2}\cdot\frac{2}{3}=\frac{1}{3}$ |
| 0 Scissors | −1, 1<br>0 | 1, −1<br>0 | 0, 0<br>0 |

(a) Player 1 never plays Scissors, so the whole Scissors row has probability zero. Player 2 never plays Rock, so the whole Rock column has probability zero. That leaves four cells. One of those four, (Paper, Paper), is a tie with utility 0. So **three terms are non-zero**.

(b) For Player 1, we use the first number in each of the three cells:

$$u_1(s_1,s_2) = \underbrace{\frac{1}{2}\cdot\frac{1}{3}}_{P(\text{Rock, Paper})}\cdot(-1) \quad + \quad \underbrace{\frac{1}{2}\cdot\frac{2}{3}}_{P(\text{Rock, Scissors})}\cdot 1 \quad + \quad \underbrace{\frac{1}{2}\cdot\frac{2}{3}}_{P(\text{Paper, Scissors})}\cdot(-1)$$
$$= -\frac{1}{6}+\frac{1}{3}-\frac{1}{3} \;=\; -\frac{1}{6}.$$

(c) For Player 2, we keep the probabilities and switch to the second number in each cell:

$$u_2(s_1,s_2) = \frac{1}{2}\cdot\frac{1}{3}\cdot 1 \quad + \quad \frac{1}{2}\cdot\frac{2}{3}\cdot(-1) \quad + \quad \frac{1}{2}\cdot\frac{2}{3}\cdot 1 \;=\; \frac{1}{6}-\frac{1}{3}+\frac{1}{3} \;=\; \frac{1}{6}.$$

The two expected utilities add up to zero, as they must in a zero-sum game.

**Solution 2: Best responses in the Prisoner's Dilemma**

We answer one fixed column at a time. In each column, we compare Player 1's two utilities (underlined).

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

(a) If Player 2 Cooperates

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

(b) If Player 2 Defects

(a) If Player 2 Cooperates, Player 1 gets $-1$ by Cooperating and $0$ by Defecting. The best response is Defect.

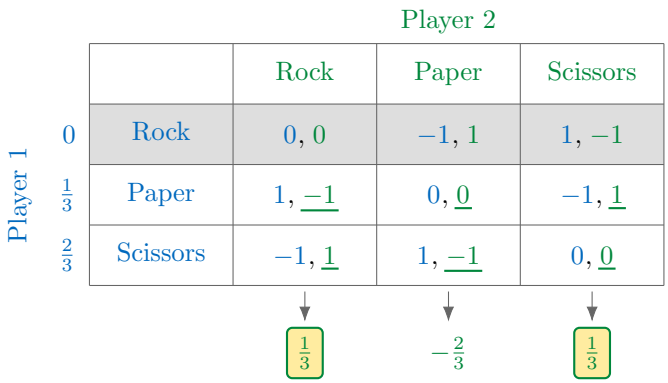
(b) If Player 2 Defects, Player 1 gets $-3$ by Cooperating and $-2$ by Defecting. The best response is again Defect.

Player 1's best response is the same whatever Player 2 does. The game is symmetric, so the same is true for Player 2.

Yet if both players Defect, each gets 2 years, when both could have had 1 year by Cooperating. That is the dilemma, and we will come back to it in lecture.

**Solution 3: When best responses tie**

We are choosing for the column player. So we use the second number in each cell, and we get one expected utility for each *column*, weighted by Player 1's probabilities.




(a) Player 1 never plays Rock, so the Rock row (shaded gray) contributes nothing, and each sum has two terms:

$$u_2(s_1, \text{Rock}) = \frac{1}{3} \cdot (-1) + \frac{2}{3} \cdot 1 = \frac{1}{3}$$
$$u_2(s_1, \text{Paper}) = \frac{1}{3} \cdot 0 + \frac{2}{3} \cdot (-1) = -\frac{2}{3}$$
$$u_2(s_1, \text{Scissors}) = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 0 = \frac{1}{3}$$

Rock and Scissors tie for the highest expected utility, so both are best responses. Paper is not.

(b) With $s_2 = \left(\frac{1}{2}, 0, \frac{1}{2}\right)$, half of the time Player 2 plays Rock, with expected utility $\frac{1}{3}$, and half of the time Scissors, also with expected utility $\frac{1}{3}$:

$$u_2(s_1, s_2) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{3}.$$

That ties the highest value, so yes, this mixed strategy is a best response too. So is every other mix of Rock and Scissors.

A strategy that puts any probability on Paper is not a best response, because it averages in the $-\frac{2}{3}$.

So Player 2's best response to $s_1$ is a whole set: Rock, Scissors, and every mixed strategy whose support contains only Rock and Scissors. We will see in lecture that this pattern is general.

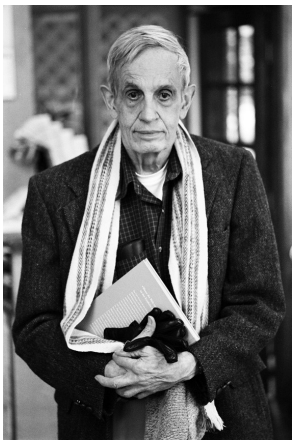
# B Game Theory: People and Applications

This appendix is background and is not needed for lecture. Numbers in brackets refer to the sources in Section B.6.

## B.1 John Nash

Game theory grew out of mathematicians' studies of games such as chess and cards. John von Neumann introduced the minimax solution for two-player zero-sum games in 1928. His 1944 book with Oskar Morgenstern, *Theory of Games and Economic Behavior*, laid the foundations for using game theory in economics [1].

John Nash (1928–2015) came to the Carnegie Institute of Technology, now Carnegie Mellon, on a George Westinghouse Scholarship. He started in chemical engineering, moved to chemistry, and then to mathematics, and when he graduated in 1948 he was given an M.S. in addition to his B.S. [2] [3].



John Nash in the 2000s. Photo: Peter Badge / Typos1, CC BY-SA 3.0, via Wikimedia Commons

His Carnegie Tech professor Richard Duffin recommended him to Princeton in a letter of just three sentences. The first two say that Nash has applied to Princeton, is nineteen years old, and is graduating from Carnegie Tech in June. The third is:

> "He is a mathematical genius." [3]

At Princeton, Nash wrote a PhD thesis of about thirty pages, *Non-Cooperative Games* (1950) [4]. Earlier analysis had centered on two-player zero-sum games. Nash formulated a solution concept for any number of players and any utilities, now called the **Nash equilibrium** [1]. He proved that every game with a finite number of players and actions has at least one equilibrium when mixed strategies are allowed [3].

Nash shared the 1994 Nobel Memorial Prize in Economic Sciences with John Harsanyi and Reinhard Selten "for their pioneering analysis of equilibria in the theory of non-cooperative games" [1]. His life is the subject of Sylvia Nasar's biography *A Beautiful Mind* (1998) and of the 2001 film of the same name [3].

## B.2 Auctions

An auction is a game: the best bid for one bidder depends on what the other bidders do.

In 1994, the U.S. government first sold radio frequencies to telecom operators in an auction. The format was invented by Paul Milgrom and Robert Wilson, partly with Preston McAfee. Over the next twenty years, FCC auctions in this format brought in more than 120 billion dollars. Milgrom and Wilson received the 2020 Nobel Memorial Prize in Economic Sciences "for improvements to auction theory and inventions of new auction formats" [5].

Auctions also run constantly online. Each time an ad is eligible to appear for a Google search, it goes through an ad auction [6].

## B.3 Security and Wildlife Protection

Where should a limited number of guards patrol when an attacker can watch for patterns? A predictable schedule can be exploited, so the answer is a randomized one: a mixed strategy.

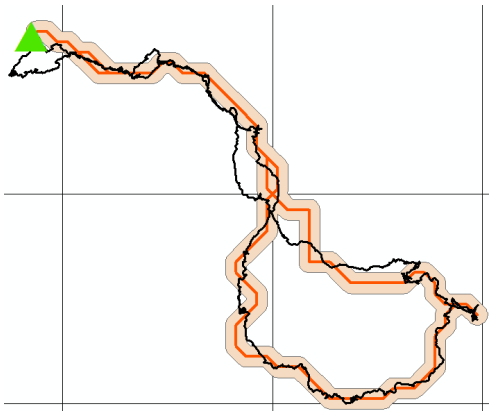
Game-theoretic systems from Milind Tambe's research group have been deployed for security at Los Angeles International Airport and for U.S. Coast Guard protection of ports [7] [8].

Fei Fang, now an associate professor in CMU's School of Computer Science [9], worked on these "security games" as a PhD student in that group. She worked on a Coast Guard system to protect the Staten Island Ferry [8], and she is the first author of the paper on deploying PAWS, the Protection Assistant for Wildlife Security [10].

PAWS helps rangers protect wildlife from poachers. Its output is a probability distribution over patrol routes, which is a mixed strategy. It was first tested in Uganda in 2014 and then went into regular use in Malaysia [10].



Fei Fang with a tiger snare found in a wildlife preserve in northeast China. Photo: Yongchao Jing, from IEEE Spectrum [8]



One daily PAWS patrol route in Malaysia, August 2015.
Figure 9 of Fang et al. [10]

## B.4 Poker

In January 2017, Libratus, an AI developed at CMU by Professor Tuomas Sandholm and PhD student Noam Brown, played 120,000 hands of two-player no-limit Texas hold'em against four of the world's best professional players at Rivers Casino in Pittsburgh, and won [11].

Poker is a game of imperfect information: the players cannot see each other's cards. The name Libratus is Latin for "balanced and forceful," where "balanced" refers to approximating a Nash equilibrium [12].



"Tuomas Sandholm (center) and Ph.D. student Noam Brown developed Libratus." Photo and caption: Carnegie Mellon University [11]

In 2019, Pluribus, developed by CMU in collaboration with Facebook AI, defeated leading professionals at six-player no-limit Texas hold'em. One of them, Darren Elias, said of it: "Its major strength is its ability to use mixed strategies" [13].

## B.5 Machine Learning

A generative adversarial network (GAN) trains two models at once: a generator, and a discriminator that estimates whether a sample came from the training data or from the generator. In the words of the original paper, this "corresponds to a minimax two-player game" [14].

AlphaZero reached a superhuman level of play in chess, shogi, and Go, starting from random play, by reinforcement learning from games of self-play [15].

## B.6 Sources


1. The Royal Swedish Academy of Sciences, press release for the 1994 prize in economic sciences.
2. John F. Nash Jr., autobiography, The Nobel Prize, 1994.
3. Princetoniana Museum, John Forbes Nash Jr. *50. The letter is transcribed in full in W. Gasarch, Computational Complexity blog, March 3, 2024.
4. John F. Nash Jr., *Non-Cooperative Games*, PhD thesis, Princeton University, 1950. Scan from the Princeton University Library (32 pages with its front matter).
5. The Royal Swedish Academy of Sciences, popular science background for the 2020 prize in economic sciences.
6. Google Ads Help, Auction.
7. F. Fang, P. Stone, and M. Tambe, When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing, IJCAI 2015. Its introduction cites the deployed systems: ARMOR at Los Angeles International Airport (Pita et al., AAMAS 2008) and PROTECT for the ports of the United States (Shieh et al., AAMAS 2012).
8. This AI Hunts Poachers, IEEE Spectrum, January 2018.
9. Fei Fang, home page.
10. F. Fang, T. H. Nguyen, R. Pickles, W. Y. Lam, G. R. Clements, B. An, A. Singh, M. Tambe, and A. Lemieux, Deploying PAWS: Field Optimization of the Protection Assistant for Wildlife Security, IAAI 2016.
11. Carnegie Mellon University News, Carnegie Mellon Artificial Intelligence Beats Top Poker Pros, January 31, 2017.
12. N. Brown and T. Sandholm, Superhuman AI for heads-up no-limit poker: Libratus beats top professionals, Science 359, 418–424, 2018.
13. Carnegie Mellon University News, Carnegie Mellon and Facebook AI Beats Professionals in Six-Player Poker, July 11, 2019.
14. I. Goodfellow et al., Generative Adversarial Networks, 2014.
15. D. Silver et al., Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm, 2017.