Game Theory for Single Interactions

CMPUT 261: Introduction to Artificial Intelligence

S&LB §3.0-3.3.2, 3.4.1

Lecture Overview

1. Recap & Logistics
2. Game Theory
3. Solution Concepts
4. Mixed Strategies
5. Minimax Strategies

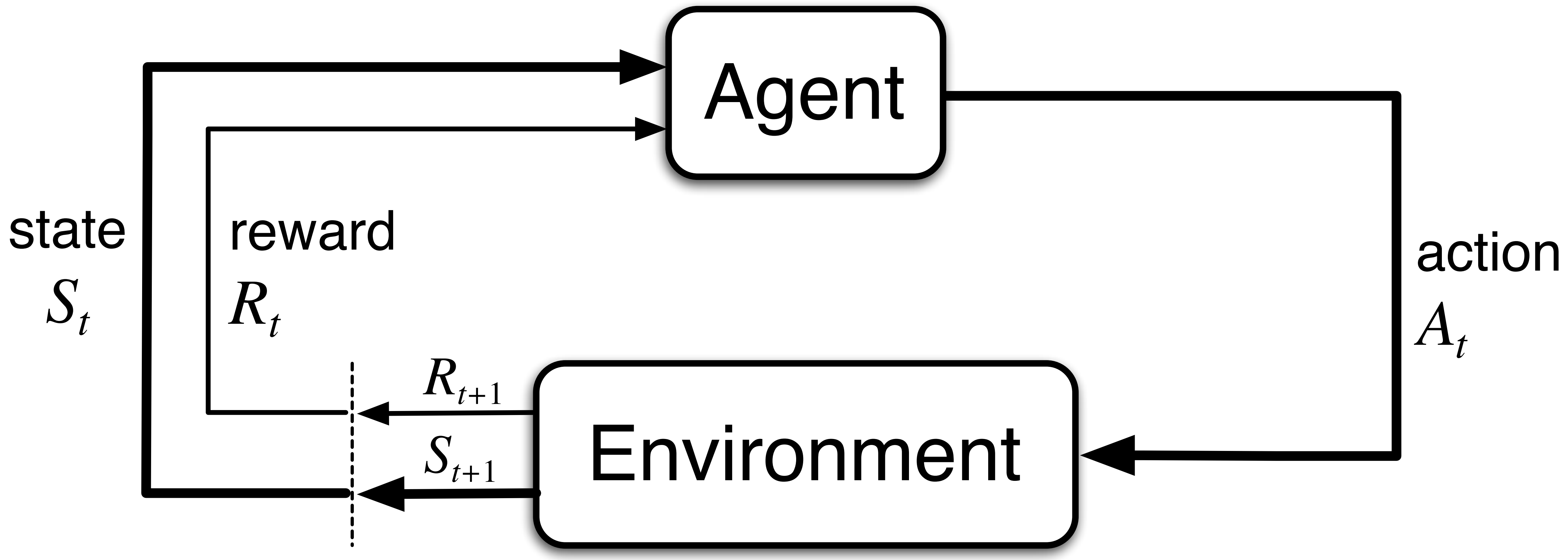
After this lecture, you should be able to:

- define best response and Nash equilibrium
- define Pareto dominance and Pareto optimality
- identify the pure strategies in a normal form game
- identify a pure strategy Nash equilibrium in a normal form game
- identify the Pareto dominant outcomes in a normal form game
- explain the difference between pure strategy and mixed strategy Nash equilibria
- define a maxmin strategy
- define a zero-sum game
- state the Minimax Theorem and explain its implications

Logistics

- **Assignment #4** is due **TODAY** at 11:59pm
  - Late submissions for 20% deduction until Thursday at 11:59pm
  - We plan to have **on-time submissions** marked before the final; no guarantees whatsoever for late submissions
- **SPOT** (formerly USRI) surveys are now available: https://p20.courseval.net/etw/ets/et.asp?nxappid=UA2&nxmid=start
  - You should have gotten an email
  - Available until **Friday** (April 14)
  - Please do fill one out for this class!
- **Final exam** is Thursday, **April 20** at **9:00am** in **CCIS-L2-200** (*not* this room!)

Recap: Reinforcement Learning



- **Reinforcement learning:** Single agents learn from **interactions** with an **environment**
- **Prediction:** Learn the value $v_\pi(s)$ of executing **policy** π from a given **state** s, or the value $q_\pi(s, a)$ of taking **action** a from state s and *then* executing π
- **Control:** Learn an optimal **policy**
  - **Action-value methods: Policy improvement** based on action value estimates
  - **Policy gradient methods:** Search **parameterized policies** directly

Game Theory

- **Game theory** is the mathematical study of interaction between multiple **rational**, self-interested agents
- **Rational** agents' preferences can be represented as maximizing the **expected value** of a **scalar utility function**
- **Self-interested:** Agents pursue only their **own preferences**
  - *Not* the same as "agents are psychopaths"! Their preferences may include the well-being of other agents.
  - Rather, the agents are **autonomous**: they decide on their own priorities independently.

How Is This AI?

- We will not be talking about algorithms for constructing agents today
- All of our material up until today has assumed **one agent** interacting with an **environment**
- As we'll see today, things are very different when the "environment" contains **other agents** with **distinct** preferences and goals
- Reasoning about **incentives** is crucial when multiple agents interact
- **Game theory** is a principled way to reason about incentives

Fun Game:
Prisoner's Dilemma

	Cooperate	Defect
Cooperate	-1,-1	-5,0
Defect	0,-5	-3,-3

Two suspects are being questioned separately by the police.

- If they both remain silent (**cooperate** -- i.e., with each other), then they will both be sentenced to **1 year** on a lesser charge
- If they both implicate each other (**defect**), then they will both receive a reduced sentence of **3 years**
- If one defects and the other cooperates, the defector is given immunity (0 years) and the cooperator serves a full sentence of **5 years**.

Play the game with someone near you. Then find a new partner and play again.

Normal Form Games

The Prisoner's Dilemma is an example of a **normal form game**.
Agents make a single decision **simultaneously**, and then receive a payoff depending on the profile of actions.

Definition: Finite, n-person normal form game

- N is a set of n ***players***, indexed by i
- $A = A_1 \times A_2 \times \cdots \times A_n$ is the set of **action profiles**
    - A_i is the **action set** for player i
- $u = (u_1, u_2, \ldots, u_n)$ is a **utility function** for each player
    - $u_i : A \to \mathbb{R}$

Utility Theory

- The expected value of a **scalar** utility function $u_i : A \to \mathbb{R}$ is sufficient to represent "rational preferences" [von Neumann & Morgenstern, 1944]
  - **Rational preferences** are those that satisfy **completeness**, **transitivity**, **substitutability**, **decomposability**, **monotonicity**, and **continuity**
  - **Action profile** determines the **outcome** in a normal form game
- **Affine invariance:** For a given set of preferences, u_i is not unique
  - $u_i'(a) = au_i(a) + b$ represents the same preferences $\forall a > 0,\ b \in \mathbb{R}$ (**why?**)

Games of Pure Cooperation and Pure Competition

- In a **zero-sum game**, players have **exactly opposed** interests:
  $u_1(a) = -u_2(a)$ for all $a \in A$ **(*)**
  - There must be **precisely two** players
- In a game of **pure cooperation**, players have **exactly the same** interests:
  $u_i(a) = u_j(a)$ for all $a \in A$ and $i, j \in N$

	Heads	Tails
Heads	1,-1	-1,1
Tails	-1,1	1,-1

Matching Pennies

	Left	Right
Left	1	-1
Right	-1	1

Which side of the road should you drive on?

General Game: Battle of the Sexes

The most interesting games are simultaneously both **cooperative and competitive**!

	Ballet	Soccer
Ballet	2, 1	0, 0
Soccer	0, 0	1, 2

Play against someone near you.

Optimal Decisions in Games

- In single-agent environments, the key notion is
  optimal decision: a decision that maximizes the agent's expected utility
- **Question:** What is the **optimal strategy** in a multiagent setting?
  - In a multiagent setting, the notion of unconditional optimal strategy is **incoherent**
  - The best strategy **depends** on the strategies of others

Solution Concepts

- From the viewpoint of an **outside observer**, can some outcomes of a game be labelled as **better** than others?
  - We have no way of saying one agent's interests are more important than another's
  - We can't even **compare** the agents' utilities to each other, because of affine invariance! We don't know what "**units**" the payoffs are being expressed in.
- Game theorists identify certain subsets of outcomes that are interesting in one sense or another. These are called **solution concepts**.

Pareto Optimality

- Sometimes, some outcome o^1 is **at least as good** for **any** agent as outcome o^2, and there is some agent who **strictly prefers** o^1 to o^2.
  - In this case, o^1 seems defensibly better than o^2

Definition: o^1 **Pareto dominates** o^2 in this case

Definition: An outcome o^* is **Pareto optimal** if no other outcome Pareto dominates it.

$\begin{bmatrix} 9 \\ 8 \\ 7 \end{bmatrix}$ Pareto-dominates $\begin{bmatrix} 8 \\ 2 \\ 7 \end{bmatrix}$

$\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ Pareto-dominates $\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$

Does $\begin{bmatrix} 9 \\ 8 \\ 7 \end{bmatrix}$ Pareto-dominate $\begin{bmatrix} 1 \\ 1 \\ 9 \end{bmatrix}$?

Out of $\left\{ \begin{bmatrix} 9 \\ 8 \\ 7 \end{bmatrix}, \cancel{\begin{bmatrix} 8 \\ 2 \\ 7 \end{bmatrix}}, \cancel{\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}}, \cancel{\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}}, \begin{bmatrix} 1 \\ 1 \\ 9 \end{bmatrix} \right\}$ which outcomes are Pareto-optimal?

Questions:

1. Can a game have more than one Pareto-optimal outcome?
2. Does every game have at least one Pareto-optimal outcome?

Best Response

- Which **actions** are better from an **individual agent's** viewpoint?
- That depends on what the other agents are doing!

Notation:

$$a_{-i} \doteq (a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$$

$$a = (a_i, a_{-i})$$

Definition: Pure Best Besponse

$$BR_i(a_{-i}) \doteq \{a_i^* \in A_i \mid u_i(a^*, a_{-i}) \geq u_i(a_i, a_{-i}) \ \ \forall a_i \in A_i\}$$

Nash Equilibrium

- Best response is not, in itself, a solution concept
  - In general, agents won't know what the other agents will do
  - But we can use it to define a solution concept
- A **Nash equilibrium** is a **stable** outcome: one where no agent regrets their actions

Definition:
An action profile $a \in A$ is a (pure strategy) **Nash equilibrium** iff

$$\forall i \in N, \ a_i \in BR_i(a_{-i})$$

Questions:

1. Can a game have **more than one** pure strategy Nash equilibrium?
2. Does every game have **at least one** pure strategy Nash equilibrium?

Nash Equilibria of Examples

The only equilibrium of Prisoner's Dilemma is also the *only* outcome that is Pareto-dominated!

	Coop.	Defect
Coop.	-1,-1	-5,0
Defect	0,-5	-3,-3

	Left	Right
Left	1	-1
Right	-1	1

	Ballet	Soccer
Ballet	2, 1	0, 0
Soccer	0, 0	1, 2

	Heads	Tails
Heads	1,-1	-1,1
Tails	-1,1	1,-1

Mixed Strategies

Definitions:

- A **strategy** s_i for agent i is any probability distribution over the set A_i, where each action a_i is played with probability $s_i(a_i)$.
  - **Pure strategy:** only a single action is played
  - **Mixed strategy:** randomize over multiple actions
- Set of i's strategies: $S_i \doteq \Delta(A_i)$

  $\Delta(X)$ is the set of distributions over elements of X
- Set of **strategy profiles:** $S = S_1 \times S_2 \times \cdots \times S_n$
- **Utility** of a mixed strategy profile:

$$u_i(s) \doteq \sum_{a \in A} u_i(a) \prod_{j \in N} s_j(a_j)$$

Best Response and Nash Equilibrium

Definition:
The set of i's **best responses** to a strategy profile $s \in S$ is

$$BR_i(s_{-i}) \doteq \{s_i^* \in S_i \mid u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i}) \;\; \forall s_i \in S_i\}$$

Definition:
A strategy profile $s \in S$ is a **Nash equilibrium** iff

$$\forall i \in N, \quad s_i \in BR_i(s_{-i})$$

- When at least one s_i is mixed, s is a **mixed strategy Nash equilibrium**

Nash's Theorem

Theorem: [Nash 1951]
Every game with a finite number of players and action profiles has at least one Nash equilibrium.

- **Pure strategy** equilibria are *not* guaranteed to exist

Interpreting
Mixed Strategy Nash Equilibrium

What does it even mean to say that agents are playing a mixed strategy Nash equilibrium?

- They truly are **sampling a distribution** in their heads, perhaps to **confuse** their opponents (e.g., soccer, other zero-sum games)
- The distribution represents the **other agents' uncertainty** about what the agent will do
- The distribution is the **empirical frequency** of actions in repeated play
- The distribution is the frequency of a pure strategy in a **population** of pure strategies (i.e., every individual plays a pure strategy)

Maxmin Strategies

What is the maximum expected utility that an agent can **guarantee** themselves?

Definition:
The **maxmin value** of a game for i is the value $\overline{v}_i$ highest value that i can **guarantee** they will receive:

$$\overline{v}_i = \max_{s_i \in S_i} \left[\min_{s_{-i} \in S_i} u_i(s_i, s_{-i}) \right]$$

Definition:
A **maxmin strategy** for i is a strategy $\overline{s}_i$ that maximizes i's worst-case payoff:

$$\overline{s}_i = \arg\max_{s_i \in S_i} \left[\min_{s_{-i} \in S_i} u_i(s_i, s_{-i}) \right]$$

Question:

1. Does a maxmin strategy always **exist**?
2. Is an agent's maxmin strategy always **unique**?
3. Why would an agent **want** to play a maxmin strategy?

Minimax Theorem

Theorem: [von Neumann, 1928]
In any **Nash equilibrium** s^* of any **finite, two-player, zero-sum game**, each player receives an expected utility v_i equal to *both* their maxmin and their minmax value.

Proof sketch:

1. Suppose that $v_i < \bar{v}_i$. But then i could guarantee a higher payoff by playing their maxmin strategy. So $v_i \geq \bar{v}_i$.
2. $-i$'s equilibrium payoff is $v_{-i} = \max_{s_{-i}} u_{-i}(s_i^*, s_{-i})$
3. Equivalently, $v_i = \min_{s_{-i}} u_i(s_i^*, s_{-i})$, since the game is zero sum.
4. So $v_i = \min_{s_{-i}} u_i(s_i^*, s_{-i}) \leq \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) = \bar{v}_i$. ∎

Because:
$u_{-i}(s) = -u_i(s)$, so
$v_i = -v_{-i}$ and
$-v_i = \max_{s_i} \left[-u_i(s_i^*, s_{-i})\right]$, and
$-v_i = -\left[\min_{s_i} u_i(s_i^*, s_{-i})\right]$.

Minimax Theorem Implications

In any **zero-sum** game:

1. Each player's maxmin value is equal to their minmax value (i.e., $\overline{v}_i = \underline{v}_i$). We call this the **value of the game**.

2. For both players, the maxmin strategies and the Nash equilibrium strategies are the **same sets**.

3. Any **maxmin strategy profile** (a profile in which both agents are playing maxmin strategies) is a Nash equilibrium. Therefore, each player gets the same payoff in every Nash equilibrium (namely, their value for the game).

Nash Equilibrium Safety: General Sum Games

- In a **general-sum** game, a **Nash equilibrium** strategy is not always a **maxmin** strategy
- **Question:** What is the **Nash equilibrium** of this game?
- **Question:** What is player 1's **maxmin strategy**? [1/15: H, 14/15: L]
  - Guarantees player 1 an expected utility of **at least 1/3**
- **Question:** Can player 1 ever **regret** playing a Nash equilibrium against a **non-equilibrium** player?

	X	Y
H	-9, 1	5, 5
L	1, 0	0, 1

Nash Equilibrium Safety: Zero-sum Games

- In a zero-sum game, every **Nash equilibrium** strategy is **also** a maxmin strategy
- **Question:** What is player 1's **maxmin value**?
- **Question:** Can player 1 ever regret playing a Nash equilibrium strategy against a **non-equilibrium** player?

	X	Y
H	-9, 9	5, -5
L	1, -1	0, 0

Summary

- Game theory studies the **interactions of rational agents**
  - Canonical representation is the **normal form game**
- Game theory studies **solution concepts** rather than optimal behaviour
  - "Optimal behaviour" is not clear-cut in multiagent settings
  - **Pareto optimal:** no agent can be made better off without making some other agent worse off
  - **Nash equilibrium:** no agent regrets their strategy given the choice of the other agents' strategies
  - **Maxmin** strategies maximize an agent's worst-case payoff
- In **zero-sum games**, maxmin strategies and Nash equilibrium are the same thing
  - It is always **safe** to play an equilibrium strategy in a zero-sum game