Game Theory for Single Interactions

CMPUT 366: Intelligent Systems

S&LB §3.0-3.3.2

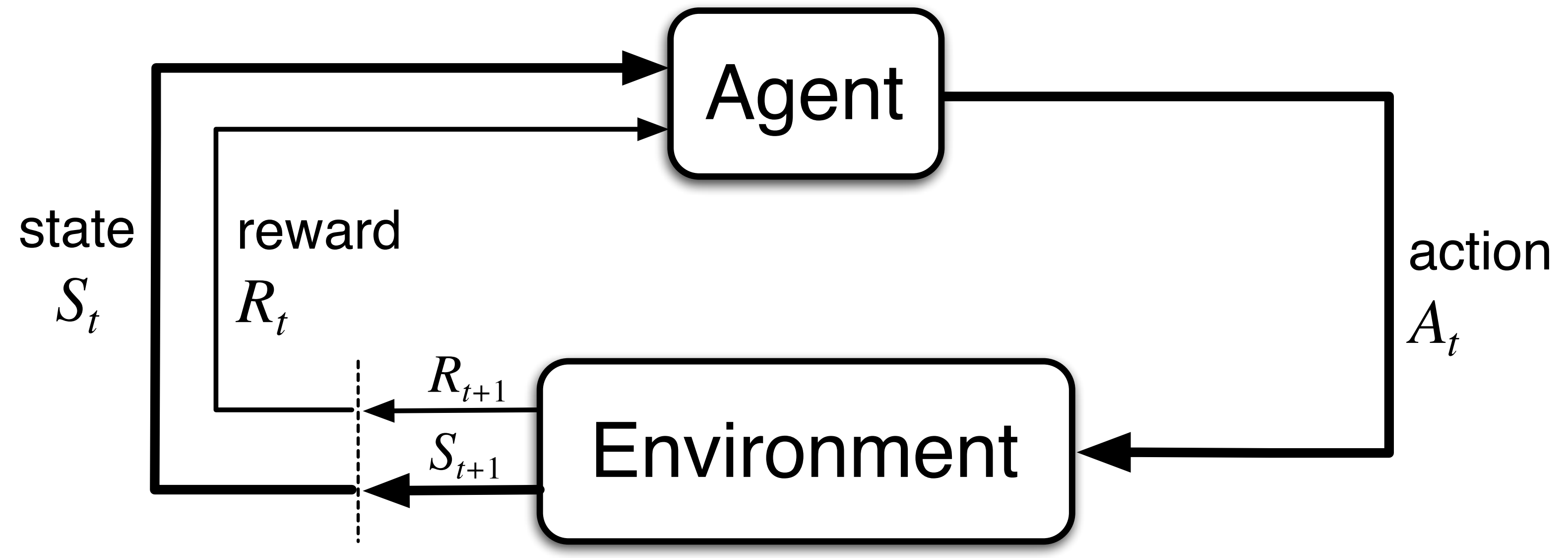
Lecture Overview

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

Labs & Assignment #4

- **Assignment #4 is released today**
  See the website under Assignments (or on the Schedule)
- Due **April 12** before **midnight**
- Today's lab is from **5:00pm to 7:50pm** in **CAB 235**
  - Not mandatory
  - You can get help from the TAs on your assignment in labs

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.

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. Play 3 times in total, against someone new each time.

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
- $A = A_1 \times A_2 \times \ldots \times A_n$ 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
  - $u_i : A \to \mathbb{R}$

Games of Pure Cooperation and Pure Competition

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

	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.
Play 3 times in total, playing against someone new each time.

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

Definition: o **Pareto dominates** o' in this case

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

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: **Best response**

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

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

Definition:

- 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: $S_i \doteq \Delta(A_i)$
- Set of **strategy profiles**: $S = S_1 \times \ldots \times S_n$
- **Utility** of a mixed strategy profile: $u_i(s) = \sum_{a \in A} u_i(a) \prod_{j \in N} s_j(a_j)$

Best Response and Nash Equilibrium

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

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

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

$$\forall i \in N, a_i \in A_i \quad s_i(a_i) > 0 \implies a_i \in BR_{-i}(s_{-i})$$

- 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)

Summary

- Game theory studies the **interactions of rational agents**
  - Canonical representation is the **normal form game**
- Game theory uses **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