Algorithmic Game Theory

Martin Schmid
Department of Applied Mathematics
Charles University in Prague

October 7, 2023

About the Course

Class

- Simultaneous and sequential decision making
- Solution concepts and optimal policies
- Practical algorithm for finding the optimal policies

Homeworks

- You will get to implement the games and algorithms!

Understand These!

(a) AlphaZero

(b) AlphaStar

(c) DeepStack

Game Theory - Reinforcement Learning

Reinforcement Learning

- Single agent settings
- Maximize reward
- Scalable practical algorithms

Game Theory

- Multi agent settings
- Analyzes agent interaction, incentives
- Optimal solution concepts
- Algorithms (historically) tabular and not scalable

Terminology

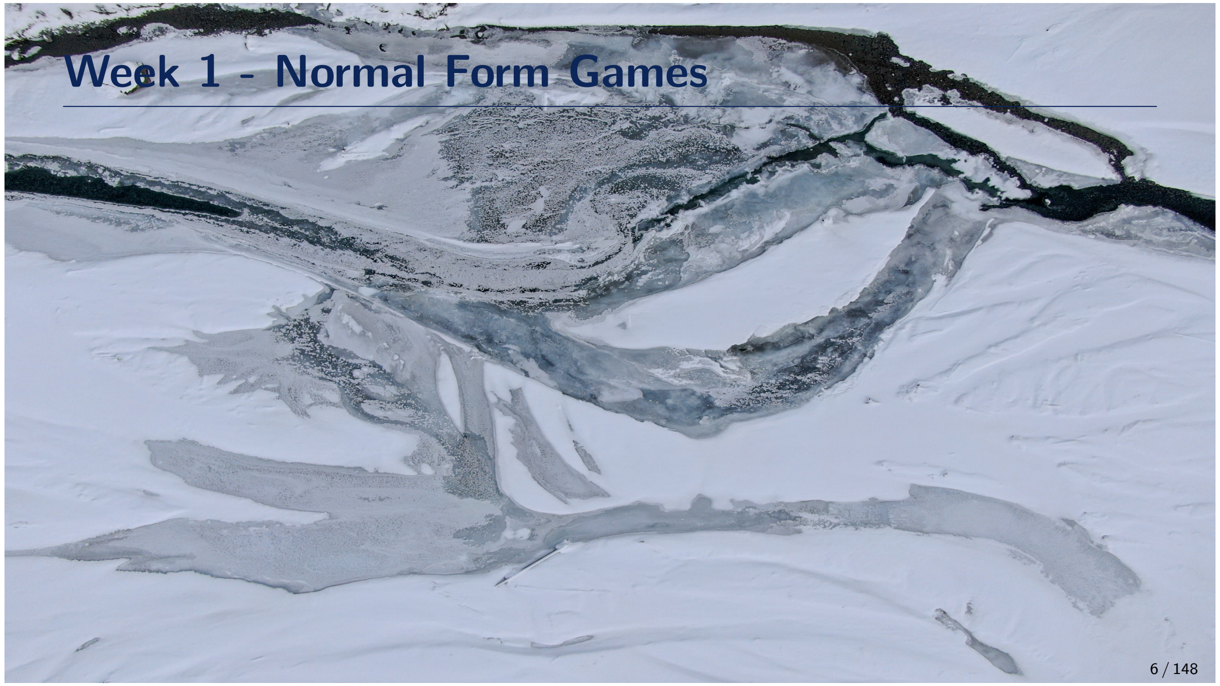
Reinforcement Learning

1. Environment
2. Agent
3. Policy
4. Reward

Game Theory

1. Game
2. Player
3. Strategy
4. Utility

Week 1 - Normal Form Games



Normal Form Games

The normal form games is a model in which each player chooses his strategy, and then all players play simultaneously. The outcome depends on the actions chosen by the players.

Definition: Normal Form Game

is a tuple $\langle N, (A_i), (u_i) \rangle$, where

- N is the **finite** set of players
- A_i is the nonempty set of actions available to the player i
- u_i is a **payoff/utility** function for the player i. Let $A = \times_{i \in N} A_i$.
$$u_i : A \to \mathbb{R}$$

Normal Form Games

- If there are only two players ($|N| = 2$) , we can conveniently described the game using a table
- Rows/columns correspond to actions of player one/two
- The cell (i, j) contains the players' payoffs $u_1(i, j)$ and $u_2(i, j)$

Normal Form Games

	Rock	Paper	Scissors
Rock	(0, 0)	(-1, 1)	(1, -1)
Paper	(1, -1)	(0, 0)	(-1, 1)
Scissors	(-1, 1)	(1, -1)	(0, 0)

Table: Rock-Paper-Scissors

	Confess	Be Quiet
Confess	(8, 8)	(0, 10)
Be Quiet	(10, 0)	(2, 2)

Table: Prisoner's dillema

Constant Sum Games

- Constant-sum game is a game for which $u_1 + u_2 = c$
- Zero-sum game is a constant-sum game for $c = 0$, so $u_1 = -u_2$
- Critical implications!

Zero Sum Games

	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**
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

Normal Form Game Strategies

Definition: Pure Strategy

$a_i \in A_i$ is player i's pure strategy. This strategy is referred to as pure, because there's no probability involved. For example, the player can always play Scissors.

Definition: Mixed Strategy

is a probability measure over the player's pure strategies. The set of player i's mixed strategies is denoted as Π_i. Given $\pi_i \in \Pi_i$, we denote the probability that the player chooses the action $a_j \in A_i$ as $\pi_i(a_j)$ Mixed strategies allow a player to probabilistically choose actions.

Normal Form Game Strategies II

Definition: Support

For a strategy π_i, support is the set of actions with non-zero probability $\{a \in A | \pi_i(a) > 0\}$.

Definition: Strategy profile

Is the set of all players' strategies, denoted as $\pi = (\pi_0, \pi_1 \cdots \pi_n)$. Finally, π_{-i} refers to all the strategies in π except π_i.

Outcome

- Given a pure strategies of all players, we can easily compute the utilities/reward. Player i's utility is $u_i(a)$.
- How to compute the outcome if the players use mixed strategy (they randomize among the pure strategies)? We simply compute the expected value given the probability measure.
- Since the players choose the actions simultaneously, the events are independent and consequently $\pi(a) = \prod \pi_i(a_i)$
- Using this fact, computing the expected value is easy
$$u_i(\pi) = \sum_{a \in A} \pi(a) u_i(a)$$

Outcome Examples

	Rock	Paper	Scissors
Rock	(0, 0)	(-1, 1)	(1, -1)
Paper	(1, -1)	(0, 0)	(-1, 1)
Scissors	(-1, 1)	(1, -1)	(0, 0)

- $\pi_1 = (0.2, 0.2, 0.6), \pi_2 = (0.2, 0.2, 0.6)$
- $\pi_1 = (0.6, 0.2, 0.2), \pi_2 = (0.2, 0.2, 0.6)$

	Confess	Be Quiet
Confess	(8, 8)	(0, 10)
Be Quiet	(10, 0)	(2, 2)

- $\pi_1 = (0.4, 0.6), \pi_2 = (0.4, 0.6)$
- $\pi_1 = (0.6, 0.4), \pi_2 = (0.4, 0.6)$

Best Response

- One of the key concepts, that you will see throughout the class
- Given the strategies π_{-i} of the opponents, the **best response** is the strategy that maximizes the utility for the player.

Definition: Best Response

Best response against a policy π_i is:

$$\arg\max_{\pi_{-i} \in \Pi_{-i}} R_{-i}(\pi_i, \pi_{-i})$$

We use $\mathbb{BR}(\pi_i)$ to denote the set of best response policies against the policy π_i.

Best Response

Note that for zero-sum games, opponent maximizing their reward is equivalent to opponent minimizing our reward.

$$\arg\max_{\pi_{-i}} R_{-i}(\pi_i, \pi_{-i}) = \arg\min_{\pi_{-i}} R_i(\pi_i, \pi_{-i})$$

As this means the player's value against any best-response strategy is unique, we denote this unique value as $BRV_i(\pi_i)$.

$$BRV_i(\pi_i) = \min_{\pi_{-i}} R_i(\pi_i, \pi_{-i}) = -\max_{\pi_{-i}} R_{-i}(\pi_i, \pi_{-i})$$

Best Response

Lemma

For any best response strategy $\pi_i \in \mathbb{BR}_i(\pi_{-i})$, all the actions in the support have the same expected value.

Lemma

The best response set $\mathbb{BR}(\pi_{-i})$ is convex.

Dominated Strategies

- Some actions can be clearly poor choises, and it makes no sense for a rational player to take.
- Strategy π_i^a **strictly dominates** π_i^b iff for any π_{-i}
$$u_i(\pi_i^a, \pi_{-i}) > u_i(\pi_i^b, \pi_{-i})$$
- Strategy π_i^a **weakly dominates** π_i^b iff for any π_{-i}
$$u_i(\pi_i^a, \pi_{-i}) \geq u_i(\pi_i^b, \pi_{-i})$$
- Strategy is **strictly/weakly** dominated if there's a strategy that strictly/weakly dominates it.
- Strategies π_i^a, π_i^b are **intransitive** iff one neither dominates nor is dominated by the other.

Examples

Can a weakly/strictly dominated strategy be a best response?

Elimination of Dominated Strategies

- A rational player does not play dominated strategy

	Left	**Center**	**Right**
Top	(13, 3)	(1, 4)	(7, 3)
Middle	(4, 1)	(3, 3)	(6, 2)
Up	(-1, 9)	(2, 8)	(8, -1)

Iterated Elimination of Dominated Strategies

- A rational player does not play dominated strategy
- Iterated elimination of dominated strategies

	Left	Center	Right
Top	(13, 3)	(1, 4)	(7, 3)
Middle	(4, 1)	(3, 3)	(6, 2)
Up	(-1, 9)	(2, 8)	(8, -1)

Iterated Elimination of Dominated Strategies

- A rational player does not play dominated strategy
- Iterated elimination of dominated strategies

	Left	**Center**	**Right**
Top	(13, 3)	(1, 4)	(7, 3)
Middle	(4, 1)	(3, 3)	(6, 2)
Up	(-1, 9)	(2, 8)	(8, -1)

Iterated Elimination of Dominated Strategies

- A rational player does not play dominated strategy
- Iterated elimination of dominated strategies

	Left	Center	Right
Top	(13, 3)	(1, 4)	(7, 3)
Middle	(4, 1)	(3, 3)	(6, 2)
Up	(-1, 9)	(2, 8)	(8, -1)

Iterated Elimination of Dominated Strategies

- A rational player does not play dominated strategy
- Iterated elimination of dominated strategies

	Left	**Center**	**Right**
Top	(13, 3)	(1, 4)	(7, 3)
Middle	(4, 1)	(3, 3)	(6, 2)
Up	(-1, 9)	(2, 8)	(8, -1)

Dominated Strategies and Best Response

Examples

Can a weakly/strictly dominated strategy that we found during the iterated elimination be a best response in the original game?

Week 1 Homework

1. Python and notebooks
2. Strategy pair evaluation for a matrix game
3. Best response calculation
4. Strategy evaluation against a best response
5. Iterated removal of dominated strategies