Repeated Games

CMPUT 355: Games, Puzzles, and Algorithms

Lecture Outline

1. Logistics & Recap
2. Dominated Strategies
3. Finitely Repeated Games
4. Indefinitely Repeated Games

Final Exam Logistics

- **Final exam: Wed April 15**
  - **CCIS L2-190** (this room)
  - **1:00pm** (*not* this time)
  - Format: like a **quiz**, but longer!
  - Cumulative: covers the **whole semester**
  - *Very slight* emphasis on post-quiz-5 (maybe one extra question)
- **Practice material:**
  - No practice final (but: finals from past offerings are available here)
  - Practice questions #1-5 will be very relevant to the final
  - Additional **practice questions #6** to be released Tue **Apr 7** (solutions **Apr 9**)

Recap: Normal Form Games

- Rock Paper Scissors is an example of a **normal form game**
- Players choose actions **simultaneously**
- Each possible **combination** of actions maps to scores for the players ("**utilities**")
- To specify a normal form game, list:
  - **Who** the players are (need not be just 2)
  - What **actions** each player has available (need not be the same per player)
  - The **utility** to each player for each combination of actions (often a table for 2-player games)
- Unlike **perfect information alternating moves**, the optimal strategy for a normal form game may require **randomization** (i.e., **mixed strategies**)
  - Players aim to maximize their **expected utility** (expectation with respect to the distribution over actions from the players' randomized strategies)

	Rock	Paper	Scissor
Rock	0,0	-1,1	1,-1
Paper	1,-1	0,0	-1,1
Scissor	-1,1	1,-1	0,0

Recap: Nash Equilibrium

- A **Nash equilibrium** is a profile of strategies in which every player simultaneously best responds to all the other players
- In the two-player case, s, t is a Nash equilibrium when:
  1. s is a best response to t; i.e., $u_r(s, t) \geq u_r(a, t)$ for all actions a
  2. t is a best response to s; i.e., $u_c(s, t) \geq u_c(s, b)$ for all actions b

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.

Dominant and Dominated Strategies

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

Questions:

1. Is a **dominated** strategy **guaranteed to exist**?
2. Can **more than one** strategy be **dominant**?

- One strategy **dominates** another if it has higher utility for **all strategies** of the other players
  - E.g., **Defect** dominates **Cooperate** in Prisoner's Dilemma
- A strategy is **dominated** if there is **any** other strategy that dominates it
  - **Cooperate** is dominated
- A strategy is **dominant** if it dominates all other strategies
  - **Defect** is dominant

Repeated Game

- Some situations are well-modelled as the **same agents** playing a normal-form game **multiple times**.
    - The normal-form game is the **stage game**; the whole game of playing the stage game repeatedly is a **repeated game.**
    - The stage game can be repeated a **finite** or an **infinite** number of times.
- Questions to consider:
    1. What do agents **observe**?
    2. What do agents **remember**?
    3. What is the agents' **utility** for the whole repeated game?

Finitely Repeated Game

Suppose that n players play a normal form game against each other $k \in \mathbb{N}$ times.

Questions:

1. Do they **observe** the other players' actions? If so, **when**?
2. Do they **remember** what happened in the previous games?
3. What is the **utility** for the whole game?
4. What are the **pure strategies**?

Fun (Repeated) Game

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

and then

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

and then

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

and then

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

and then

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

- Play the **Prisoner's Dilemma** five times in a row against the **same person**
    - Total payoff is the sum of stage-game payoffs

Properties of Finitely Repeated Games

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

and then

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

and then

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

and then

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

and then

	c	d
C	-1,-1	-4,0
D	0,-4	-3,-3

- Playing an **equilibrium of the stage game** at every stage is an equilibrium of the repeated game (**why?**)
    - Instance of a **stationary** strategy
- In general, pure strategies can depend on the **previous history** (**why?**)
- **Question:** When the normal form game has a **dominant strategy**, what can we say about the **equilibrium** of the finitely repeated game?

Infinitely Repeated Game

Suppose that n players play a normal form game against each other **infinitely many** times.

Questions:

1. Do they **remember** what happened in the previous games?
2. What is the **utility** for the whole game?
3. What are the **pure strategies**?

Payoffs in
Infinitely Repeated Games

- **Question:** What are the **payoffs** in an infinitely repeated game?
  - We cannot take the **sum of payoffs** in an infinitely repeated game (**why not?**)
  - We cannot put the overall utility on the **terminal nodes**, because there **aren't any**

Discounted Reward

Definition:
Given an infinite sequence of payoffs $r_i^{(1)}, r_i^{(2)}, \ldots$ for player i, and a discount factor $0 \leq \beta \leq 1$, the **future discounted reward** of i is

$$\sum_{t=1}^{\infty} \beta^{t-1} r_i^{(t)}.$$

- Interpretations:
  1. Agent is **impatient**: cares more about rewards that they will receive earlier than rewards they have to wait for.
  2. Agent cares equally about all rewards, but at any given round the game will **stop** with probability $1 - \beta$.
- The two interpretations have **identical implications** for analyzing the game.

Definitions: Feasible & Enforceable

Definition:
A payoff profile $r = (r_1, \ldots, r_n)$ is **feasible** if it is a convex combination of game outcomes.

That is, if there exist $\{q_a \mid a \in A\}$ with $\sum_{a \in A} q_a = 1$ such that $r_i = \sum_{a \in A} q_a u_i(a)$.

Definition:
Let v_i be i's **minmax value** in $G = (N, A, u)$.

Then a payoff profile $r = (r_1, \ldots, r_n)$ is **enforceable** if $r_i \geq v_i$ for all $i \in N$.

- A payoff vector r is **enforceable** (on i) if the other agents working together can **ensure** that i's utility is no greater than r_i.

Folk Theorem

Theorem:
Consider any n-player normal form game G and payoff profile $r = (r_1, \ldots, r_n)$.

1. If r is the payoff profile for any Nash equilibrium of the infinitely repeated G, then r is **enforceable**.
2. If r is both **feasible** and **enforceable**, then r is the payoff profile for **some** Nash equilibrium of the infinitely repeated G.

Proof sketch: ("Grim Trigger" equilibrium)

1. Construct a sequence of action profiles with the desired average
2. Everyone cycles through the action profiles in order
3. If any player fails to play "their part", then everyone **punishes** that player (by playing minimax) forever

Summary

- A **repeated game** is one in which agents play the same normal form game (the **stage game**) multiple times.
- **Finitely repeated:** Doesn't change anything when there is a **dominant strategy**
- **Infinitely repeated:** Life gets more complicated
  - Payoff to the game: either **average** or **discounted** reward
  - Pure strategies map from **entire previous history** to action
- **Folk theorem** characterizes which **payoff profiles** can arise in any equilibrium
  - All profiles that are both **enforceable** and **feasible.**