Part 7

Game theory (part I)

References

- **Multiagent Systems**
  by Yoav Shoham and Kevin Leyton-Brown.
  Cambridge University Press.
  Chapter 3.
  http://www.masfoundations.org/index.html
- **Introduction to Game Theory**
  by Dana Nau, University of Maryland.
  http://www.cs.umd.edu/users/nau/game-theory

What is game theory

Game theory is about interactions among agents that are self-interested:

- We use “agent” and “player” synonymously.

Self-interested:

- Each agent has its own description of what states are desirable.
- This is modelled using utility functions: maps each state of the world to a real number.
- They represent how much an agent likes that state.

Example: TCP Users

Internet traffic is governed by the TCP protocol.

TCP's *backoff* mechanism:

- If there is congestion, reduce the rate at which you send packets until congestion subsides.

Suppose that:

- You're trying to finish an important project.
- It's extremely important for you to have a fast connection.
- Only one other person is using the Internet.
- That person wants a fast connection just as much as you do.

You each have 2 possible actions:

- C: use a *correct* implementation.
- D: use a *defective* implementation that won't back off.

Action profiles and their payoffs

Definition

An **action profile** is a choice of action for each agent.

Action profiles are represented using tuples.

For example:

- (C, C) (both use C): the average packet delay is 1 ms.
- (D, D) (both use D): average delay is 3 ms (router overhead).
- (C, D) or (D, C) (one uses D, the other uses C):
  - D user's delay is 0 ms.
  - C user's delay is 4 ms.

Payoff matrix

Let us construct the **payoff matrix** of our game:

- Your options are the rows.
- The other agent's options are the columns.
- Each cell is an action profile:
  - The first number in the cell is your payoff or utility (we use those terms synonymously).
- The second number in each cell is the other agent's payoff.

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Some questions

Examples of the kinds of questions game theory attempts to answer:

- Which option should you use: C or D?
- Does it depend on what you think the other person will do?
- What kind of behavior can the network operator expect?
- Would any two users behave the same?
- Will this change if users can communicate with each other beforehand?
- Under what changes to the delays would the users' decisions still be the same?
- How would you behave if you knew you would face this situation repeatedly with the same person?

Some game-theoretic answers

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Suppose the only consequences are the ones in the payoff matrix:

- No other kinds of interactions between the two agents.
- No trouble from the network operator.

Suppose each user cares only about maximizing his/her own payoff.

- No guilt feelings. They do not care about the other agent's utility.

Suppose each user knows the other feels the same way.

Then they will both use D.

Some game-theoretic answers

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Allowing them to communicate beforehand won't change the outcome.

Repeat any fixed number of times gives the same outcome.

If the number of times is unbounded, they might use C instead.

Games in normal form

Definition

A (finite n-person) **normal-form game** is a tuple (N, A, u), where:

- $N = \{1, 2, \ldots, n\}$ is a finite set of agents or players.
- $A = A_1 \times \cdots \times A_n$, where each A_i is a finite set of actions available for player i.
- For each agent i, a utility (or payoff) function $u_i : A \mapsto \mathbb{R}$.

Most other game representations can be reduced to normal form.

A natural way to represent a normal-form game is with an n-dimensional payoff (or utility) matrix that shows every agent's utility for every action profile.

The prisoner's dilemma

The TCP user's game is more commonly called the Prisoner's Dilemma:

- Scenario: two prisoners are in separate rooms.
- For each prisoner, the police have enough evidence for a 1 year prison sentence.

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

- They want to get enough evidence for a 4 year prison sentence.
- They tell each prisoner:
  "If you testify against the other prisoner, your sentence will be reduced by 1 year."
- C = Cooperate (with the other prisoner): refuse to testify.
- D = Defect: testify against the other prisoner.
- If both prisoners cooperate, then they both stay in prison for 1 year.
- If both prisoners defect, then they both stay in prison for $4 - 1 = 3$ years.
- If one defects and the other cooperates, then the cooperator stays in prison for 4 years, while the defector goes free.

The prisoner's dilemma

The payoff matrix that we used:

	C	D
C	−1, −1	−4, 0
D	0, −4	−3, −3

The payoff matrix that is usually used:

	C	D
C	3, 3	0, 5
D	5, 0	1, 1

The prisoner's dilemma

The exact numbers are not important, as long as the following conditions hold:

$$c > a > d > b$$

$$a > \frac{(b + c)}{2}$$

	C	D
C	a, a	b, c
D	c, b	d, d

More generally

Under standard utility theory, games are insensitive to any positive affine transformation of the payoffs.

Replace each payoff x_i by $cx_i + d$, where c and d are constants and $c > 0$.

	C	D
C	x_1, x_2	x_3, x_4
D	x_5, x_6	x_7, x_8

This is because every positive affine transformation of the payoffs correspond to the same set of **rational preferences**.

Preferences

Game-theoretic utilities are based on preferences.

Suppose an agent can choose among:

- outcomes (A, B, etc.), and
- lotteries (situations with uncertain outcomes).

Lottery $L = [p : A; 1 - p : B]$:

- Probability p of getting A.
- Probability $1 - p$ of getting B.
- $0 \leq p \leq 1$.

Notation:

- $A \succ B$ (A is preferred to B).
- $A \sim B$ (indiference between A and B).
- $A \succsim B$ ($A \succ B$ or $A \sim B$).

Rational preferences

Idea: the preferences of a rational agent must obey some constraints.

Agent's choices are based on rational preferences. Agent's behavior is describable as maximization of expected utility.

Constraints:

- **Orderability** (sometimes called **Completeness**):
  $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
- **Transitivity:**
  $(A \succsim B) \wedge (B \succsim C) \Rightarrow (A \succsim C)$
- **Continuity:**
  $A \succ B \succ C \Rightarrow \exists p[p : A; 1 - p : C] \sim B$
- **Substitutability** (sometimes called **Independence**):
  $A \sim B \Rightarrow [p : A; 1 - p : C] \sim [p : B; 1 - p : C]$
- **Monotonicity:**
  $A \succ B \Rightarrow (p > q \Leftrightarrow [p : A; 1 - p : B] \succ [q : A; 1 - q : B])$

Rational preferences

What happens if the constraints are violated?

Example: intransitive preferences (*money pump*):

- If $B \gtrsim C$, then an agent who has C would trade C plus some money to get B.
- If $A \gtrsim B$, then an agent who has B would trade B plus some money to get A.
- If $C \succ A$, then an agent who has A would trade Aplus some money to get C.

Such an agent can be induced to give away all its money.

Violating the constraints leads to self-evident irrationality.

Utility functions

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944)

Given preferences satisfying the constraints, there exists a real-valued function u such that:

$$u(A) \geq u(B) \Leftrightarrow A \succsim B$$
$$u([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i u(S_i)$$

u is called a **utility function.**

MEU principle:

- If an agent's choices are based on rational preferences, then the agent's behavior is describable as maximization of expected utility.

An agent can maximize the expected utility without ever representing or manipulating utilities and probabilities:

- E.g., a lookup table to play tic-tac-toe perfectly.

Utility scales

Preferences are invariant with respect to positive affine transformations.

Let:

$$u'(x) = k_1 u(x) + k_2 \qquad \text{where } k_1 > 0$$

Then u' models the same set of preferences that u does.

Normalized utilities: define u such that $u_{\max} = 1$ and $u_{\min} = 0$.

Common-payoff games

A **common-payoff** game is one in which for every action profile, all agents have the same payoff.

	Left	Right
Left	1, 1	0, 0
Right	0, 0	1, 1

Also called a **pure coordination** game or a **team game**.

Need to coordinate on an action that is maximally beneficial to all.

Common-payoff games

Example: Which Side of the Road?

- 2 people driving in a country with no traffic rules.
- Coming at each other.
- Independently decide to stay left or right.

Need to coordinate your action with the action of the other driver.

How to accomplish this?

Mechanism design:

- Change the rules of the game to give each agent an incentive to choose a desired outcome.

Zero-sum games

Zero-sum games are purely competitive games.

Constant-sum game:

- For every action profile, the sum of the payoffs is the same.
- There is a constant c such that for every action profile $(a_1, \ldots, a_n)$, $u_1(a_1, \ldots, a_n) + \cdots + u_n(a_1, \ldots, a_n) = c$.

Any constant-sum game can be transformed into an equivalent game in which the sum of the payoffs is always 0:

- Just subtract $\frac{c}{n}$ from every payoff.

Thus, constant-sum games are usually called zero-sum games.

Examples

Matching pennies:

- Two agents, each has a penny.
- Each agent independently chooses to display Heads or Tails.
- If same, agent 1 gets both pennies.
- Otherwise, agent 2 gets both pennies.

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

Rock, paper, scissors (rochambeau):

- 3-action generalization of matching pennies.
- If both choose same, no winner.
- Otherwise, paper beats rock, rock beats scissors, scissors beats paper.

	Rock	Paper	Scissors
Rock	0,0	-1,1	1,-1
Paper	1,-1	0,0	-1,1
Scissors	-1,1	1,-1	0,0

Nonzero-sum games

A game is nonzero-sum if $u_1(a_1, \ldots, a_n) + \cdots + u_n(a_1, \ldots, a_n)$ is different for different action profiles.

- E.g., the prisoner's dilemma.

Nonzero-sum games include aspects of both coordination and competition.

Battle of the sexes:

- Two agents need to coordinate their actions, but they have different preferences.
- Original scenario:
  - Wife prefers opera.
  - Husband prefers football.
- Another scenario:
  - Two nations must act together to deal with an international crisis.
  - They prefer different solutions.

	Opera	Football
Opera	$2, 1$	$0, 0$
Football	$0, 0$	$1, 2$

Symmetric games

A game is symmetric if:

- Both agents have the same set of actions.
- A action's payoff is independent of which agent uses it.

For a 2×2 symmetric game, the payoff matrix looks like this:

	C	D
C	a, a	b, c
D	c, b	d, d

Most of the games we have seen are symmetric.

Symmetric games

	C	D
C	a	b
D	c	d

In symmetric games, we only need to show u_1 in the matrix:

- $u_1(r, r) = u_2(r, r) = a =$ the payoff r gets against itself.
- $u_1(r, r') = u_2(r', r) = b =$ the payoff r gets against r'.
- $u_1(r', r) = u_2(r, r') = c =$ the payoff r' gets against r.
- $u_1(r', r') = u_2(r', r') = d =$ the payoff r' gets against itself.

Symmetric games

As originally stated, the battle of the sexes is not symmetric.

	Opera	Football
Opera	2, 1	0, 0
Football	0, 0	1, 2

But by renaming the strategies, we can transform it into an equivalent game that is symmetric:

	Give	Take
Take	2,1	0,0
Give	0,0	1,2

Most games cannot be transformed like that.

Strategies in normal-form games

Pure strategy: select a single action and play it.

- Each row or column of a payoff matrix represents both an action and a pure strategy.

Pure-strategy profile: a choice of pure strategy for each agent.

Mixed strategy: randomize over the set of available actions according to some probability distribution.

- $s_i(a_i) =$ probability that action a_i will be played under mixed strategy s_i.

The **support** of s_i is the set $\mathsf{support}(s_i) = \{a_i \in A_i \mid s_i(a_i) > 0\}$.

A pure strategy is a special case of a mixed strategy (the support is a singleton).

Fully mixed strategy: every action has probability > 0.

Expected utility

A payoff matrix only gives payoffs for pure-strategy profiles.

Generalization to mixed strategies uses **expected utility.**

First, calculate probability of each outcome, given the strategy profile (involves all agents).

Then calculate average payoff for agent i, weighted by the probabilities.

For a strategy profile $s = (s_1, \ldots, s_n)$, the expected utility is:

$$u_i(s) = \sum_{(a_1,\ldots,a_n)\in A} u_i(a_1, \ldots, a_n) \prod_{j=1}^{n} s_j(a_j)$$

where $s_j(a_j)$ is the probability that an action a_j will be played under mixed strategy s_j