Lecture Title

CMPUT 658: Modelling Human Strategic Behaviour

S&LB §12.1-12.2.2

Lecture Outline

1. Coalitional Game Theory Setting
2. Shapley Value
3. The Core

Coalitional Game Theory / Cooperative Game Theory

- The game theory we've thought about up until now is often called **noncooperative game theory**
  - Aims at modeling how agents choose their **actions**
- Coalitional game theory abstracts away from actions entirely
  - Sets of agents (**coalitions**) cooperate in some unspecified way to obtain **value**
  - Different coalitions can obtain **different** total value
  - Aims to analyze the **division** of the value among the coalition members
  - Often the focus is on the "Grand Coalition" that contains all agents

Formal Setting

Definition:

A **coalitional game** (with transferable utility) is a pair (N, v), where:

- N is a set of **agents**
- $v : 2^N \to \mathbb{R}$ is the **characteristic function**, which maps each subset of agents to the quantity of **value** that the subset can obtain.

- If a given coalition $S \subseteq N$ **forms**, then it obtains $v(S)$
  - This value must then be **distributed** among the coalition members
- Typically we want to ask two main questions:

1. Which coalitions will form?
2. How should the coalition divide its payoff among its members?

Example: Voting Game

Voting game:

A parliament has 4 parties (a, b, c, d), with 45, 25, 15, and 15 members. They are voting on a spending bill, which requires 51 votes to pass. If the bill passes, then the parties that voted for the bill will each control some portion of the spending; if the bill fails, then no party gets any money to spend.

- We can model the game as follows:

$$f(a) = 45, f(b) = 25, f(c) = 15, f(d) = 15$$

$$v(S) = \begin{cases} 100 & \text{if } \sum_{j \in S} f(j) \geq 51 \\ 0 & \text{otherwise.} \end{cases}$$

- More generally:
  - Some set $W \subseteq 2^N$ of **winning coalitions**
  - $v(S) = \begin{cases} 1 & \text{if } S \in W \\ 0 & \text{otherwise.} \end{cases}$

Example: Airport Game

Airport Game:

Several neighbouring cities need airport capacity. The cost of an airport depends on the largest aircraft that its runway can support. If a new regional airport is built, then the cities will have to share its cost; otherwise, each city will have to build its own airport.

- N is the set of cities
- $c(x)$ is the cost of building an airport that can support aircraft of size x
- $x(j)$ is the largest runway required by city j

$$v(S) = \left(\sum_{j \in S} c(x(j)) \right) - c\left(\max_{k \in S} x(k) \right)$$

Dividing Coalition Value

Definition: A **payoff vector** is a vector $x \in \mathbb{R}^N$ that satisfies the following:

- $\forall j \in N : x_j \geq 0$
- $\sum_{j \in N} x_j \leq v(N)$

- A payoff vector is the formal object that describes how the grand coalition's payoffs will be divided
- A (semi-)**value function** $\psi(N, v) \to \mathbb{R}^N$ maps a coalitional game to an payoff vector

Axiom: Efficiency

Axiom: A value function ψ satisfies **efficiency** if

$$\sum_{i \in N} \psi_i(N, v) = v(N)$$

- We obviously don't want to fail to allocate any of the value

Fairness Axiom: Symmetry

Definition: Two agents $i, j \in N$ are **interchangeable** if

$$\forall S \subseteq N \backslash \{i, j\} : v(S \cup \{i\}) = v(S \cup \{j\})$$

Axiom: A value function ψ satisfies **symmetry** if

$$\psi_i(N, V) = \psi_j(N, V)$$

for all interchangeble agents i, j.

- Two agents who contribute the same to every coalition should be treated the same way.

Fairness Axiom: Dummy Player

Definition: An agent $i \in N$ is a **dummy player** if

$$v(S \cup \{i\}) = v(S) + v(\{i\}) \quad \forall S \in N \backslash \{i\}$$

Axiom: If $i \in N$ is a dummy player, then

$$\psi_i(N, v) = v(\{i\}) \, .$$

- Any player that brings the same additive amount to any coalition should just receive the amount that they bring

Fairness(?) Axiom: Additivity

Axiom: A value function ψ satisfies **additivity** if, for all $i \in N$ and v_1, v_2,

$$\psi_i(N, v_1 + v_2) = \psi_i(N, v_1) + \psi_i(N, v_2),$$

where $(v_1 + v_2)(S) = v_1(S) + v_2(S)$ for all $S \subseteq N$.

- A coalition that plays in two different games should receive the same amount as they would have received in a single game that has the same total payoff as the two games
- **Question:** Why is this desirable?

The Shapley Value

Theorem: The Shapley value is the unique value function that satisfies the Efficiency, Dummy player, Symmetry, and Additivity axioms.

Definition: The **Shapley value** is the value function defined by

$$\psi_i(N, v) = \frac{1}{|N|!} \sum_{S \subseteq N \setminus \{i\}} |S|!(|N| - |S| - 1)! \left[v(S \cup \{i\}) - v(S) \right]$$

Shapley Value Walkthrough

Definition: The **Shapley value** is the value function defined by

$$\psi_i(N, v) = \frac{1}{|N|!} \sum_{S \subseteq N \setminus \{i\}} |S|!(|N| - |S| - 1)! \left[v(S \cup \{i\}) - v(S) \right]$$

- The Shapley value imagines each agent joining the grand coalition in some order
- The difference $\left[v(S \cup \{i\}) - v(S) \right]$ agent i's **marginal contribution** to the coalition assuming that the agents of S joined **before** them
- Marginal contribution depends on the **specific order** that the agents joined
- The Shapley value **averages** over all possible joining orders
- There are $|S|!(|N| - |S| - 1)!$ orderings in which S joined before i (**why?**)
- There are $|N|!$ possible joining orders

Which Coalitions Form?

- Some coalitions may not be "worth it" to their members
- Examples:
  1. If $v(S) < \sum_{j \in S} v(\{j\})$, then we would not expect coalition S to form (**why?**)
  2. If $v(S) < v(S \setminus \{i\})$ for some $i \in S$ (why?)
- These examples are driven by the **differences in payments** that a subcoalition might receive
- If there is no way to divide up the payoffs of the coalition that is satisfactory, then the coalition is **unstable**; we shouldn't expect it to form (or persist)

The Core

Definition: A payoff vector $x \in \mathbb{R}^N$ is in **the core** of a coalitional game (N, v) iff

$$\forall S \subseteq N : \sum_{j \in S} x_j \geq v(S)$$

- Intuitively, a payoff vector is in the core if it is a way to keep the grand coalition stable
- A subcoalition S for which this condition fails is called a **blocking coalition**
- **Question:** Is the core guaranteed to be **non-empty**?
- **Question:** Is the core guaranteed to contain only a **single** payoff vector?

Summary

- Coalitional game theory models a situation where a cooperating group of agents must **divide up the payoffs** that they have collectively created
  - Abstracts away from all details of **how** the payoffs are created!
- The **Shapley value** is a rule for computing a payoff division
  - Uniquely satisfies Efficiency, Symmetry, Dummy player, Additivity
- The **core** is the set of payoff vectors for which no block coalition exists
  - i.e., no coalition that would be better off leaving the grand coalition and earning payoffs on their own