Multiagent Systems

The Shapley Value

Tomáš Kroupa

2020

Department of Computer Science
Faculty of Electrical Engineering
Czech Technical University in Prague

How to divide the prey fairly?

Four animal friends hunt buffaloes together. The pig eats the dead prey and the owl can track down buffaloes. Only the Komodo dragon and the tiger can kill a buffalo.

$$v(A) = \begin{cases} 2 & A = N, 234 \\ 1 & A = 134, 34 \\ 0 & \text{otherwise} \end{cases}$$

What are basic principles of fair allocation?

$$v(A) = \begin{cases} 2 & A = N, 234 \\ 1 & A = 134, 34 \\ 0 & \text{otherwise} \end{cases}$$

The same reward for the same working contribution

The dragon and the tiger should get an equal portion!

He who does not work, neither shall he eat

The pig should not get any portion of the prey!

Banzhaf: Weighted voting doesn't work

Nassau County Board's voting system in 1967

Each town has the number of votes based on its population:

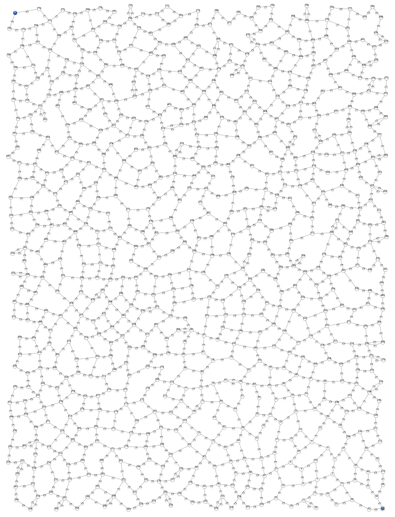
Town i	*Votes* w_i
1	31
2	31
3	28
4	21
5	2
6	2

$$v(A) = \begin{cases} 1 & \sum_{i \in A} w_i \geq 58 \\ 0 & \text{otherwise} \end{cases}$$

Neither 5 nor 6 can overturn any decision!

Voting power is the ability of a legislator to cast a decisive vote.

Why is the Shapley value important for computer science?



Quantify the effect of individual components on the performance of the entire system.

More applications

- Google Analytics
- Explainable AI algorithms
- Computational biology
- Who has the power in EU

The Shapley value: Axioms and Computation

What is the value of a game?

Definition

Let Γ be the set of all games over N. Value is a mapping

$$\varphi\colon \Gamma \to \mathbb{R}^n, \qquad \varphi = (\varphi_1, \dots, \varphi_n).$$

A value φ is

- efficient if $\varphi_1(v) + \dots + \varphi_n(v) = v(N)$ for each $v \in \Gamma$,
- additive if $\varphi(u + v) = \varphi(u) + \varphi(v)$ for all $u, v \in \Gamma$.

Formulating basic axioms of fairness

The same reward for the same working contribution

Value φ is symmetric if the following implication holds for each game v, players $i, j \in N$, and each coalition $A \subseteq N \setminus ij$:

$$v(A \cup i) = v(A \cup j) \qquad \Longrightarrow \qquad \varphi_i(v) = \varphi_j(v).$$

He who does not work, neither shall he eat

Value φ satisfies the null player property, if the following implication holds for each game v, each $i \in N$, and all $A \subseteq N \setminus i$:

$$v(A \cup i) = v(A) \qquad \Longrightarrow \qquad \varphi_i(v) = 0$$

The axioms determine the unique value for all games!

Theorem (Shapley, 1953)

There is a unique value $\varphi^S : \Gamma \to \mathbb{R}^n$, which is efficient, additive, symmetric, and satisfies the null player property.

The Shapley value of player $i \in N$ is

$$\varphi_i^S(v) = \sum_{A \subseteq N \setminus i} \frac{|A|!(n - |A| - 1)!}{n!} \cdot \big(v(A \cup i) - v(A)\big)$$

How to represent a game in a computer?

Hardly. An n-player game is determined by 2^n numbers.

The linear space of all n-player games Γ

Every $v \in \Gamma$ can be viewed as a real vector in $\mathbb{R}^{2^n}$. For example:

$$(\underbrace{v(\emptyset)}_{0}, v(1), v(2), v(3), v(12), v(13), v(23), v(123)) \in \mathbb{R}^8$$

$$\Downarrow$$

Γ is a real vector space of dimension $2^n - 1$

The axioms determine the unique value: Why?

For each coalition $A \neq \emptyset$, define a game

$$u_A(B) := \begin{cases} 1 & A \subseteq B \\ 0 & \text{otherwise} \end{cases}$$

Fact 1

$\{u_A \mid \emptyset \neq A \subseteq N\}$ is a basis of Γ

Fact 2

If a value φ is efficient, symmetric, and has the null player property, then, for every $\alpha \in \mathbb{R}$ and each coalition $A \neq \emptyset$,

$$\varphi_i(\alpha \cdot u_A) = \begin{cases} \frac{\alpha}{|A|} & i \in A \\ 0 & \text{otherwise} \end{cases}$$

How to divide the prey fairly – the Shapley value

$$v(A) = \begin{cases} 2 & A = N, 234 \\ 1 & A = 134, 34 \\ 0 & \text{otherwise} \end{cases}$$

The Shapley values of animals

- $\varphi_3^S(v) = \varphi_4^S(v)$ Symmetry
- $\varphi_1^S(v) = 0$ Null player property
- $\varphi_2^S(v) + 2\varphi_3^S(v) = 2$ Efficiency

$$\varphi_2^S(v) = \frac{1}{12} + \frac{1}{4} = \frac{1}{3} \qquad \Longrightarrow \qquad \varphi^S(v) = \left(0, \frac{2}{6}, \frac{5}{6}, \frac{5}{6}\right)$$

The Shapley values as expected values (1)

$$\varphi_i^S(v) = \sum_{A \subseteq N \setminus i} \frac{1}{n\binom{n-1}{|A|}} \cdot \underbrace{(v(A \cup i) - v(A))}_{\text{marginal contribution of } i \text{ to } A}$$

The probability $\frac{1}{n\binom{n-1}{|A|}}$ is determined in two stages:

1. Player i randomly selects the size of a coalition to enter
2. Coalition A of this size is then randomly chosen

The Shapley values as expected values (2)

$$\varphi_i^S(v) = \sum_{\pi \in \Pi} \frac{1}{n!} \cdot \underbrace{\left(v(A_i^\pi \cup i) - v(A_i^\pi)\right)}_{\text{marginal vector } x_i^\pi}$$

- Players arrive at random order given by permutation π
- Each permutation π of players has probability $\frac{1}{n!}$

Estimation of the Shapley value

Algorithm

Input: Game v over n players and a selected player i

1. Pick the size of the random sample $m \ll n!$
2. Sample with replacement permutations $(\pi_1, \ldots, \pi_m)$ with uniform probability $\frac{1}{n!}$
3. Estimate the Shapley value of player i by

$$\sum_{k=1}^{m} \frac{1}{m} \cdot x_i^{\pi_k}$$

The algorithm is polynomial, provided that worth $v(A)$ of each coalition A can be calculated in polynomial time.

The Shapley value of simple games

How to determine the voting power in simple games?

- Simple game v captures collective binary decision-making
- Player $i \in N$ is pivotal to a coalition $A \subseteq N \setminus i$ if

$$v(A \cup i) = 1 \qquad \text{and} \qquad v(A) = 0$$

Definition

The Shapley–Shubik index of player i in a simple game v is

$$\varphi_i^S(v) = \sum_{\substack{A \subseteq N \setminus i \\ i \text{ pivotal to } A}} \frac{1}{n\binom{n-1}{|A|}}$$

The voting power in a simple majority game

Simple Majority Game

$N = \{1, \dots, n\}$

$$v(A) = \begin{cases} 1 & |A| > \frac{n}{2}, \\ 0 & \text{otherwise}, \end{cases} \qquad A \subseteq N.$$

- $\varphi_i^S(v) = \varphi_j^S(v)$ for each $i, j \in N$ — Symmetry
- $\varphi_1^S(v) + \cdots + \varphi_n^S(v) = 1$ — Efficiency

$$\boldsymbol{\varphi}^S(v) = \left(\tfrac{1}{n}, \dots, \tfrac{1}{n}\right)$$

The voting power in the UN Security Council

UN Security Council

- 5 *permanent* and 10 *non-permanent* members
- A binary decision is approved by all the permanent members and ≥ 4 non-permanent members

$$v(A) = \begin{cases} 1 & \text{if } A \supseteq \{1, \ldots, 5\} \text{ and } |A| \geq 9, \\ 0 & \text{otherwise.} \end{cases}$$

- $i \in \{6, \ldots, 15\} \implies \varphi_i^S(v) = \binom{9}{3} \cdot \frac{8! \cdot 6!}{15!} \approx 0.0019$
- $j \in \{1, \ldots, 5\} \implies$ symmetry and efficiency give

$$\varphi_j^S(v) = \tfrac{1}{5}(1 - 10 \cdot \varphi_i^S(v)) \approx 0.1963$$

The Shapley–Shubik index from random order

The probability that player i casts the decisive vote in an order given by randomly chosen permutation $\pi \in \Pi$:

$$\varphi_i^S(v) = \frac{1}{n!} \cdot |\{\pi \in \Pi \mid i \text{ is pivotal to } A_i^\pi\}|$$

Example (Weighted majority voting)

Each player $i \in \{1, 2, 3\}$ has i votes. The pivotal players are:

$$12\boxed{3} \qquad 1\boxed{3}2 \qquad 21\boxed{3} \qquad 2\boxed{3}1 \qquad 3\boxed{1}2 \qquad 3\boxed{2}1$$

$$\varphi^S(v) = \left(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{4}{6}\right)$$

Beyond the Shapley value

Is there an alternative to the Shapley-Shubik index?

- A swing for player i is a coalition in which i is pivotal
- Define

$$\begin{aligned} s_i(v) &= |\{A \subseteq N \setminus i \mid A \text{ is a swing for } i\}| \\ &= |\{A \subseteq N \setminus i \mid v(A \cup i) - v(A) = 1\}| \end{aligned}$$

Definition

The Banzhaf index of player i in a simple game v is

$$\varphi_i^B(v) = \frac{1}{2^{n-1}} \cdot s_i(v)$$

The Banzhaf index – example

Example (Weighted majority voting)

Each player $i \in \{1,2,3\}$ has i votes. The swings for players are:

Player i	*Coalitions*	$s_i(v)$
1	$\emptyset, 2, \boxed{3}, 23$	1
2	$\emptyset, 1, \boxed{3}, 13$	1
3	$\emptyset, \boxed{1}, \boxed{2}, \boxed{12}$	3

$$\varphi^B(v) = \left(\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{3}{4}\right)$$

Note that $\varphi_1^B(v) + \varphi_2^B(v) + \varphi_3^B(v) \neq 1$

Normalizing the Banzhaf index

Definition

The normalized Banzhaf index of player i in a simple game v is

$$\beta_i(v) = \frac{s_i(v)}{s_1(v) + \cdots + s_n(v)}$$

The two Banzhaf indices preserve the players' power ratios:

$$\beta_i(v) = \frac{2^{n-1}}{s_1(v) + \cdots + s_n(v)} \cdot \varphi_i^B(v)$$

Example

UN Security Council – The old and the new voting system

O 11 members, approval by at least 7 votes

N 15 members, approval by at least 9 votes

Shapley–Shubik indices

O $\varphi_1^S(v) = 0.1974$, $\varphi_6^S(v) = 0.0022$ 90 : 1

N $\varphi_1^S(v) = 0.1963$, $\varphi_6^S(v) = 0.0019$ 100 : 1

Normalized Banzhaf indices

O $\beta_1(v) = \frac{19}{105}$, $\beta_6(v) = \frac{1}{63}$ 11 : 1

N $\beta_1(v) = \frac{106}{635}$, $\beta_6(v) = \frac{21}{1270}$ 10 : 1

Comparison of power indices for simple games

Property/Index	*Shapley-Shubik*	*Banzhaf*	*normalized Banzhaf*
Symmetry	✓	✓	✓
Null player property	✓	✓	✓
Efficiency	✓	–	✓