Computational Game Theory

The Shapley Value

Tomáš Kroupa

2021

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 the 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!

John 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:

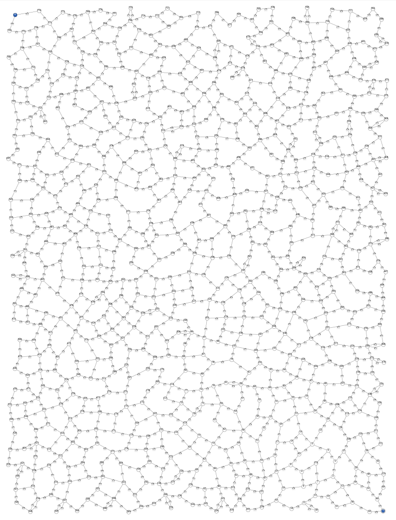
Town i	*Votes* w_i
1	31
2	31
3	28
4	21
5	2
6	2

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

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
- Centralities in networks

The Shapley value: From Axioms to the Formula

How to associate a unique allocation with any game?

Definition

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

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

The vector $\boldsymbol{\varphi}(v) = (\varphi_1(v), \dots, \varphi_n(v))$

- provides a unique outcome of any game v
- can be interpreted as the allocation to the players decided by an external arbitrator or by an authority

Any value φ must reflect basic axioms of fairness.

Formulating basic axioms of fairness (1)

A player $i \in N$ is called a null player in a game v if

$$v(A \cup i) = v(A), \qquad \text{for all } A \subseteq N \setminus i.$$

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 and each player $i \in N$:

$$i \text{ is a null player} \qquad \Longrightarrow \qquad \varphi_i(v) = 0$$

Formulating basic axioms of fairness (2)

Players $i, j \in N$ are symmetric in a game v if

$$v(A \cup i) = v(A \cup j), \qquad \text{for each coalition } A \subseteq N \setminus ij.$$

The same reward for the same working contribution

Value φ is symmetric if the following implication holds for each game v and all players $i, j \in N$:

$$i \text{ and } j \text{ are symmetric} \qquad \Longrightarrow \qquad \varphi_i(v) = \varphi_j(v).$$

The axioms determine the Shapley value for all games!

A value φ is

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

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

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

The Shapley value as expected value

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

Then $\sum_{A \subseteq N \setminus i} p_i(A) = 1$ and the probability distribution p_i corresponds to the following random scheme:

- Players in N are randomly ordered
- $p_i(A)$ is the probability that i is preceded by A:

$$A,\ i,\ N \setminus (A \cup i)$$

Example: The Shapley value for 3-player games

$N = \{1, 2, 3\}$

The Shapley value of player 2 is determined as follows:

Coalition A	**Permutations**	$p_2(A)$
$\emptyset$	213, 231	$^1/_3$
1	123	$^1/_6$
3	321	$^1/_6$
13	132, 312	$^1/_3$

$$\varphi_2(v) = \tfrac{1}{3}v(2) + \tfrac{1}{6}\big(v(12) - v(1)\big) + \tfrac{1}{6}\big(v(23) - v(3)\big) + \tfrac{1}{3}\big(v(123) - v(13)\big)$$

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 axioms determine the Shapley 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 the vector space 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 $A \neq \emptyset$,

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

Alternative axioms for the Shapley value

The motivation: Get rid of the the additivity property!

Theorem (Young, 1985)

The Shapley value φ^S is the unique value φ, which is efficient, symmetric, and satisfies marginality:

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

for all games $v, w \in \Gamma$, any player $i \in N$ and any $A \subseteq N \setminus i$

An alternative Shapley value formula

The Shapley value is the expected value of marginal vectors $\boldsymbol{x}^\pi$:

$$\varphi_i^S(v) = \frac{1}{n!} \sum_{\pi \in \Pi} \underbrace{v(A_i^\pi \cup i) - v(A_i^\pi)}_{x_i^\pi} \qquad \forall i \in N$$

Permutation	**Marginal vector $\boldsymbol{x}^{\pi_k}$**
123	$(v(1), v(12) - v(1), v(N) - v(12))$
132	$(v(1), v(N) - v(13), v(13) - v(1))$
213	$(v(12) - v(2), v(2), v(N) - v(12))$
231	$(v(N) - v(23), v(2), v(23) - v(2))$
312	$(v(13) - v(3), v(N) - v(13), v(3))$
321	$(v(N) - v(23), v(23) - v(3), v(3))$
$\boldsymbol{\varphi}^S(v)$	$\frac{1}{6}(\boldsymbol{x}^{\pi_1} + \cdots + \boldsymbol{x}^{\pi_6})$

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 Shapley value of simple games

How to determine the voting power in a simple game v?

Player $i \in N$ is pivotal to $A \subseteq N \setminus i$ if $v(A \cup i) - v(A) = 1$.

The Shapley value of simple games

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

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

$$= \frac{1}{n!} \cdot \left|\{\pi \in \Pi \mid i \text{ is pivotal to } A_i^\pi\}\right|$$

Example 1: The Shapley–Shubik index from random order

Example (Weighted majority voting)

$N = \{1, 2, 3\}$

Each player $i \in N$ has i votes:

$$v(A) = \begin{cases} 0 & A = \emptyset, 1, 2, 3, 12 \\ 1 & A = 13, 23, 123 \end{cases}$$

For each permutation the pivotal players are:

$12\mathbf{3} \quad 1\mathbf{3}2 \quad 21\mathbf{3} \quad 2\mathbf{3}1 \quad 3\mathbf{1}2 \quad 3\mathbf{2}1$

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

Example 2: The voting power in a simple majority game

Simple Majority Game

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

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

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

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

Example 2: 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

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

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

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

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

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

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 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	✓	–	✓