Relaxations of symmetry and the weighted Shapley values

André Casajus\textsuperscript{a,b}

\textsuperscript{a}HHL Leipzig Graduate School of Management, Jahnallee 59, 04109 Leipzig, Germany
\textsuperscript{b}Dr. Hops Craft Beer Bar, Eichendorffstr. 7, 04277 Leipzig, Germany

Abstract

We revisit Kalai and Samet’s (Int J Game Theory 16, 1987, 205–222) first characterization of the class of weighted Shapley values. While keeping efficiency, additivity, and the null player property from the original characterization of the symmetric Shapley value, they replace symmetry with positivity and partnership consistency. The latter two properties, however, are neither implied by nor related to symmetry. We suggest relaxations of symmetry that together with efficiency, additivity, and the null player property characterize classes of weighted Shapley values. For example, weak sign symmetry requires the payoffs of mutually dependent players to have the same sign. Mutually dependent players are symmetric players whose marginal contributions to coalitions containing neither of them are zero.

Keywords: TU game, weighted Shapley values, sign symmetry, mutual dependence, weak sign symmetry, superweak sign symmetry

2010 MSC: 91A12

JEL: C71, D60

1. Introduction

The (symmetric) Shapley value (Shapley, 1953b) probably is the most eminent one-point solution concept for cooperative games with transferable utility (TU games or simply games). In order to account for asymmetries among players beyond the game, Shapley (1953a) himself suggests a weighted version of his value, where these asymmetries are modelled by strictly positive weights for the players, the positively weighted Shapley values. Later on, Kalai and Samet (1987) extend the class of positively weighted Shapley values by considering weight systems that allow for zero weights of the players.

There exists a number of axiomatic foundations for classes of weighted Shapley values. Two types of axiomatizations can be distinguished: axiomatizations with exogenous weights and axiomatizations with endogenous weights. In axiomatizations with exogenous weights, the players’ weights are explicitly given and the properties used may involve the weights (see, e.g., Hart and Mas-Colell, 1989; Calvo and Santos, 2000). In contrast, axiomatizations with exogenous weights aim at characterizing whole classes of weighted Shapley values, i.e., the
properties used do not involve the weights, but which implicitly are given by the solutions themselves (see, e.g., Kalai and Samet [1987]; Hart and Mas-Colell [1989]; Chun [1991]; Nowak and Radzik [1995]; Casajus [2018b]).

The modern version of the original characterization of the Shapley value involves efficiency, additivity, the null player property, and symmetry. Kalai and Samet [1987] replace symmetry with positivity and partnership consistency. The latter two properties, however, are neither implied by nor related to symmetry. We suggest relaxations of symmetry that together with efficiency, additivity, and the null player property characterize classes of weighted Shapley values.

Symmetry requires equal payoffs for equally productive players, i.e., players whose marginal contributions to coalitions containing neither of them are equal. Weak sign symmetry relaxes symmetry by strengthening its hypothesis and weakening its implication. It requires equal signs of the payoffs of mutually dependent players. Mutually dependent players are symmetric players who are only jointly productive, i.e., their marginal contributions to coalitions containing neither of them are zero (Nowak and Radzik [1995]). Superweak sign symmetry relaxes weak sign symmetry by only requiring the payoffs of mutually dependent payoffs not to have opposite signs. As our main result, we show that the class of positively weighted Shapley values is characterized by efficiency, additivity, the null player property, and weak sign symmetry, while the full class of weighted Shapley values is characterized by efficiency, additivity, the null player property, and superweak sign symmetry (Theorem 3).

The remainder of this paper is organized as follows. In Section 2, we provide basic definitions and notation. In Section 3, we present Kalai and Samet’s characterizations of the weighted Shapley values. In Section 4, we introduce our relaxations of symmetry. In Section 5, we present our characterizations of the weighted Shapley values. Some remarks conclude this paper.

2. Basic definitions and notation

A (TU) game for the finite player set \( N \) is given by a coalition function \( v : 2^N \to \mathbb{R}, v(\emptyset) = 0 \), where \( 2^N \) denotes the power set of \( N \). Subsets of \( N \) are called coalitions; \( v(S) \) is called the worth of coalition \( S \). The set of all games is denoted by \( \mathcal{V} \). Let \( n := |N| \).

For \( v, w \in \mathcal{V} \) and \( \alpha \in \mathbb{R} \), the coalition functions \( v + w \in \mathcal{V} \) and \( \alpha \cdot v \in \mathcal{V} \) are given by \( (v + w)(S) = v(S) + w(S) \) and \( (\alpha \cdot v)(S) = \alpha \cdot v(S) \) for all \( S \subseteq N \). The game \( 0 \in \mathcal{V} \) given by \( 0(S) = 0 \) for all \( S \subseteq N \) is called the null game. For \( T \subseteq N, T \neq \emptyset \), the game \( u_T \in \mathcal{V} \) given by \( u_T(S) = 1 \) if \( T \subseteq S \) and \( u_T(S) = 0 \) otherwise for all \( S \subseteq N \) is called a unanimity game. A game \( v \in \mathcal{V} \) is called monotonic, if \( v(S) \leq v(T) \) for all \( S, T \subseteq N \) such that \( S \subseteq T \). Any \( v \in \mathcal{V} \) can be uniquely represented by unanimity games. In particular, we have

\[
v = \sum_{T \subseteq N : T \neq \emptyset} \lambda_T(v) \cdot u_T,
\]

where the coefficients \( \lambda_T(v) \) are known as the Harsanyi dividends (Harsanyi [1959]) and can
be determined recursively by

$$\lambda_T (v) := v (T) - \sum_{S \subseteq T : S \neq \emptyset} \lambda_S (v).$$

(2)

Player $i \in N$ is called a null player in $v \in V$ if $v (S \cup \{ i \}) = v (S)$ for all $S \subseteq N \setminus \{ i \}$; players $i, j \in N$ are called symmetric in $v \in V$ if $v (S \cup \{ i \}) = v (S \cup \{ j \})$ for all $S \subseteq N \setminus \{ i, j \}$.

A solution/value for $V$ is an operator $\varphi : V \to \mathbb{R}^N$ that assigns to any $v \in V$ and $i \in N$ a payoff $\varphi_i (v)$. The (symmetric) Shapley value \cite{Shapley1953} for $V$, $\text{Sh}$, is given by

$$\text{Sh}_i (v) := \sum_{T \subseteq N : i \in T} |T|^{-1} \cdot \lambda_T (v)$$

for all $v \in V$ and $i \in N$.

For $w \in \mathbb{R}_{++}^N$, the $w$-weighted Shapley value for $V$, $\text{Sh}^w$, is given by

$$\text{Sh}_i^w (v) := \sum_{T \subseteq N : i \in T} \sum_{\ell \in T} w_i \cdot \lambda_T (v)$$

for all $v \in V$ and $i \in N$.

The solutions $\text{Sh}^w$, $w \in \mathbb{R}_{++}^N$, are called the positively weighted Shapley values \cite{Shapley1953a}.

A weight system for $N$ is a pair $(w, \pi)$ consisting of a weight vector $w \in \mathbb{R}_{++}^N$ and a level mapping $\pi : N \to \{ 1, 2, \ldots, n \}$ with $\pi (N) = \{ 1, 2, \ldots, \# \pi \}$, where $\# \pi := |\pi (N)|$. Let $\Omega$ denote the set of all weight systems for $N$. For $(w, \pi) \in \Omega$, the $(w, \pi)$-weighted Shapley value for $V$, $\text{Sh}^{(w, \pi)}$, is given by

$$\text{Sh}_i^{(w, \pi)} (v) := \sum_{T \subseteq N : i \in T (\pi)} \sum_{\ell \in T (\pi)} w_i \cdot \lambda_T (v)$$

for all $v \in V$ and $i \in N$,

(5)

where

$$T (\pi) := \{ i \in T \mid \pi (i) \leq \pi (\ell) \text{ for all } \ell \in T \}.$$

The solutions $\text{Sh}^{(w, \pi)}$, $(w, \pi) \in \Omega$ are called the weighted Shapley values \cite{Kalai1987}. If $\# \pi = 1$, then $\text{Sh}^{(w, \pi)} = \text{Sh}^w$.

In the following, we make use of some standard properties of solutions.

**Efficiency, E.** For all $v \in V$, we have $\sum_{\ell \in N} \varphi_\ell (v) = v (N)$.

**Additivity, A.** For all $v, z \in V$, we have $\varphi_\ell (v + z) = \varphi_\ell (v) + \varphi_\ell (z)$ for all $\ell \in N$.

**Null player, N.** For all $v \in V$ and $i \in N$ such that $i$ is a null player in $v$, we have $\varphi_i (v) = 0$.

**Symmetry, S.** For all $v \in V$ and $i, j \in N$ such that $i$ and $j$ are symmetric in $v$, we have $\varphi_i (v) = \varphi_j (v)$.

**Null game, NG.** For all $i \in N$, we have $\varphi_i (0) = 0$. 

3
3. Kalai and Samet’s characterizations of the weighted Shapley values

The modern version of Shapley original characterization of the symmetric Shapley value uses efficiency, additivity, the null player property, and symmetry. Based on this characterization, Kalai and Samet (1987, Theorems 2 and 3) suggest first characterizations of the classes of (positively) weighted Shapley values for fixed player sets. They replace symmetry with partnership consistency and either positivity or strict positivity.

**Positivity, P.** For all \( v \in \mathcal{V} \) such that \( v \) is monotonic, we have \( \varphi_\ell(v) \geq 0 \) for all \( \ell \in N \).

**Strict positivity, P*.** For all \( v \in \mathcal{V} \) such that \( v \) is monotonic and there are no null players in \( v \), we have \( \varphi_\ell(v) > 0 \) for all \( \ell \in N \).

A coalition \( P \subseteq N \) is called a **partnership** in \( v \in \mathcal{V} \), if \( v(S) = v(S \cup T) \) for all \( S \subseteq N \setminus P \) and \( T \subseteq P \).

**Partnership consistency, PC.** For all \( P \subseteq N \), \( P \neq \emptyset \) and \( v \in \mathcal{V} \) such that \( P \) is a partnership in \( v \), we have

\[
\varphi_\ell(v) = \varphi_\ell \left( \sum_{\ell' \in P} \varphi_{\ell'}(v) \cdot u_P \right) \quad \text{for all } \ell \in P.
\]

Although partnership consistency looks a bit technical, the intuition behind it is sound. The players in a partnership \( P \) behave essentially as one player in the game \( v \) because any proper subcoalition is unproductive. This one player takes her payoff from the game and then the players in \( P \) bargain on the distribution of this payoff. In a sense, however, partnership consistency rather directly invokes the idea of “weightedness” of a solution.

**Theorem 1 (Kalai and Samet, 1987).** (i) If a solution \( \varphi \) for \( \mathcal{V} \) satisfies efficiency (\( E \)), additivity (\( A \)), the null player property (\( N \)), partnership consistency (\( PC \)), and positivity (\( P \)), then there exists a weight system \( \omega \in \Omega \) such that \( \varphi = Sh^\omega \).

(ii) If a solution \( \varphi \) for \( \mathcal{V} \) satisfies efficiency (\( E \)), additivity (\( A \)), the null player property (\( N \)), partnership consistency (\( PC \)), and strict positivity (\( P^* \)), then there exists a weight vector \( w \in \mathbb{R}^N_+ \) such that \( \varphi = Sh^w \).

4. Relaxations of symmetry

It is clear that the weighted Shapley values (except the symmetric Shapley value) fail symmetry. Recently, Casajus (2018a) suggests a relaxation of symmetry called sign symmetry. Recall the sign function \( \text{sign}: \mathbb{R} \to \{-1, 0, 1\} \) given by \( \text{sign}(x) = 1 \) for \( x > 0 \), \( \text{sign}(0) = 0 \), and \( \text{sign}(x) = -1 \) for \( x < 0 \).

**Sign symmetry, \( S^- \).** For all \( i, j \in N \) and \( v \in \mathcal{V} \) such that \( i \) and \( j \) are symmetric in \( v \), we have

\[
\text{sign}(\varphi_i(v)) = \text{sign}(\varphi_j(v)).
\]

\(^1\)In the following, we will refrain from mentioning this exception.
Sign symmetry is a qualitative version of symmetry that relaxes symmetry in the obvious sense. Instead of equating payoffs in general, it just fixes a common reference point, the zero utility.

One easily checks that the weighted Shapley values also fail sign symmetry. Indeed, symmetry can be replaced with sign symmetry in Young’s (1985) characterization of the Shapley value by efficiency, symmetry, and strong monotonicity or marginality or (Casajus 2018a, Theorem 3) and in Shapley’s characterization of his solution.

Proposition 2. The Shapley value is the unique solution for $\mathbb{V}$ that satisfies efficiency ($E$), additivity ($A$), the null player property ($N$), and sign symmetry ($S^-$).

Proof. Since $S^-$ is weaker than $S$, it suffices to show that $E$, $A$, $N$, and $S^-$ imply $S$. Let $i, j \in N$ be symmetric in $v \in \mathbb{V}$. Set $\alpha = (\varphi_i(v) + \varphi_j(v))/2$ and consider the game $z := v - \alpha \cdot (u_{ij} + u_{ji})$. By $A$ and $N$, we have $\varphi_\ell (z) = \varphi_\ell (v)$ for all $\ell \in N \setminus \{i, j\}$. Hence, $E$ implies $\varphi_i (z) + \varphi_j (z) = 0$. Since $i$ and $j$ are symmetric in $v$, $S^-$ entails $\varphi_i (z) = \varphi_j (z) = 0$. Hence, we obtain $\varphi_k (v) = \varphi_k (\alpha \cdot u_{ij}) + \varphi_k (\alpha \cdot u_{ji}) = \alpha$ for $k \in \{i, j\}$, which concludes the proof.

Instead, as we will show later on, the weighted Shapley values satisfy further relaxations of symmetry. Two players $i, j \in N$, $i \neq j$ are called mutually dependent in $v$, if $v(S \cup \{i\}) = v(S) = v(S \cup \{j\})$ for all $S \subseteq N \setminus \{i, j\}$ (Nowak and Radzik 1995), i.e., if they are only jointly productive. Expressed more technically, players $i$ and $j$ are mutually dependent in $v$, if $\lambda_T (v) = 0$ for all $T \subseteq N$, $T \neq \emptyset$ such that $|T \cap \{i, j\}| = 1$. Note that $P \subseteq N$, $P \neq \emptyset$ is a partnership in a game if and only if any two players in $P$ are mutually dependent in this game.

Weak sign symmetry, $S^-$. For all $i, j \in N$ and $v \in \mathbb{V}$ such that $i$ and $j$ are mutually dependent in $v$, we have

$$\text{sign} (\varphi_i (v)) = \text{sign} (\varphi_j (v)).$$

Mutually dependent players are symmetric but not vice versa. Hence, weak sign symmetry relaxes sign symmetry by strengthening its hypothesis.

Superweak sign symmetry, $S^\approx$. For all $i, j \in N$ and $v \in \mathbb{V}$ such that $i$ and $j$ are mutually dependent in $v$, we have that

$$\text{sign} (\varphi_i (v)) > 0 \quad \text{implies} \quad \text{sign} (\varphi_j (v)) \geq 0.$$

Since players $i$ and $j$ are interchangeable in the definition, the contraposition of its implication entails that sign ($\varphi_i (v)$) < 0 implies sign ($\varphi_j (v)$) ≤ 0. Superweak sign symmetry relaxes weak sign symmetry by relaxing its implication. Instead of requiring equal signs, it only rules out opposite signs.

---

2Strong monotonicity: For all $v, z \in \mathbb{V}$ and $i \in N$ such that $v(S \cup \{i\}) - v(S) \geq z(S \cup \{i\}) - z(S)$ for all $S \subseteq N \setminus \{i\}$, we have $\varphi_i (v) \geq \varphi_i (z)$. Strong monotonicity implies marginality: For all $v, z \in \mathbb{V}$ and $i \in N$ such that $v(S \cup \{i\}) - v(S) = z(S \cup \{i\}) - z(S)$ for all $S \subseteq N \setminus \{i\}$, we have $\varphi_i (v) = \varphi_i (z)$. 


5. Characterizations of the weighted Shapley values

In this section, we introduce new characterizations of the classes of weighted Shapley values based on Shapley’s characterization. Replacing symmetry with superweak sign symmetry gives a characterization of the full class of weighted Shapley values. Using weak sign symmetry instead results in a characterization of the class of positively weighted Shapley values.

These characterizations of the weighted Shapley values pinpoint the relation between the symmetric Shapley value, the positively weighted Shapley values, and the weighted Shapley values to one axiom—symmetry or sign symmetry versus weak sign symmetry versus superweak sign symmetry. Moreover, these axioms become increasingly weaker, which increasingly broadens the class of solutions that satisfy all four axioms. This is not the case for the characterizations by Kalai and Samet (1987). Symmetry does not imply partnership consistency, positivity, or strict positivity. Moreover, strict positivity does not imply positivity.

Theorem 3. (i) If a solution $\varphi$ for $V$ satisfies efficiency ($E$), additivity ($A$), the null player property ($N$), and superweak sign symmetry ($S^\subseteq$), then there exists a weight system $(w, \pi) \in \Omega$ such that $\varphi = \text{Sh}^{(w, \pi)}$.

(ii) If a solution $\varphi$ for $V$ satisfies efficiency ($E$), additivity ($A$), the null player property ($N$), and weak sign symmetry ($S^=$), then there exists a weight vector $w \in \mathbb{R}^N_{++}$ such that $\varphi = \text{Sh}^w$.

Proof. We know that the weighted Shapley values satisfy $E$, $A$, and $N$. Let $i, j \in N$ be mutually dependent in $v \in V$. For $w \in \mathbb{R}^N_{++}$, we obtain

$$\text{Sh}^w_k(v) = w_k \cdot \sum_{T \subseteq N \setminus \{i, j\} \subseteq T} \frac{\lambda_T(v)}{\sum_{\ell \in T} w_\ell}$$

for $k \in \{i, j\}$, which entails that $\text{Sh}^w$ meets $S^\subseteq$. By Kalai and Samet (1987) Theorem 4, any weighted Shapley value can be approximated continuously by positively weighted Shapley values. Since the latter satisfy $S^=$, the former obey $S^\subseteq$.

Let the solution $\varphi$ satisfy $E$, $A$, $N$ and $S^\subseteq$. First, we show that $\varphi$ obeys $PC$. Let $P \subseteq N$, $P \neq \emptyset$ be a partnership in $v \in V$ and consider the game $z \in V$ given by

$$z := v - \left(\sum_{\ell \in P} \varphi_\ell(v)\right) \cdot u_P.$$

By $A$ and $N$, we have $\varphi_\ell(z) = \varphi_\ell(v)$ for all $\ell \in N \setminus \{i, j\}$. Hence, $E$ implies $\sum_{\ell \in P} \varphi_\ell(z) = 0$. Since any two players in $P$ are mutually dependent in $w$, $S^\subseteq$ implies $\varphi_\ell(z) = 0$ for all $\ell \in P$. Finally, $A$ entails

$$\varphi_\ell(v) = \varphi_\ell \left(\left(\sum_{\ell \in P} \varphi_\ell(v)\right) \cdot u_P\right)$$

for all $\ell \in P$. 

6
which proves the claim.

Careful inspection of the proof of Kalai and Samet (1987, Theorem 2) reveals that $P$ is essentially used to show two properties of $\varphi$, which can be shown using $S^\equiv$ instead of $P$. Since $\varphi$ meets $PC$, this establishes claim (i).

(a) For all $T \subseteq N$ and $i \in N$ such that $\varphi_i(u_T) \neq 0$, we have $\varphi_i(u_T) > 0$. Suppose $\varphi_i(u_T) < 0$. Since any two players in $T$ are mutually dependent in $u_T$, $S^\equiv$ entails $\varphi_\ell(u_T) \leq 0$ for all $\ell \in P \setminus \{i\}$. By $N$, we have $\varphi_\ell(u_T) = 0$ for all $\ell \in N \setminus T$. Hence, $\sum_{\ell \in N} \varphi_\ell(u_T) < 0$, which contradicts $E$.

(b) For all $T \subseteq N$, $i \in N$, and $\rho \in \mathbb{R}$, we have $\varphi_i(\rho \cdot u_T) = \rho \cdot \varphi_i(u_T)$. (*) By $A$, the claim holds for all $\rho \in \mathbb{Q}$. By $N$, we have $\varphi_\ell(\rho \cdot u_T) = 0$ for all $\ell \in N \setminus T$. Arguments similar to those for (a) show $\varphi_\ell(\rho \cdot u_T) \geq 0$ for all $\ell \in N$ and $\rho \in \mathbb{R}_{++}$. Together with $A$, this extends (*) to all $\rho \in \mathbb{R}$.

Let now the solution $\varphi$ satisfy $E$, $A$, $N$ and $S^\equiv$. Since $S^\equiv$ implies $S^=\equiv$, part (i) entails $\varphi = Sh^{(u, \pi)}$ for some $(w, \pi) \in \Omega$. Since any two players are mutually dependent in $u_N$, $S^\equiv$ and $E$ imply $\varphi_\ell(u_N) > 0$ for all $\ell \in N$. Hence, we have $\#_\pi = 1$, i.e., $\varphi = Sh^w$, which proves claim (ii).

6. Concluding remarks

We have shown that relaxing symmetry in Shapley’s characterization can either lead to another characterization of the symmetric Shapley value (Proposition 2) or to characterizations of the (positively) weighted Shapley values (Theorem 3). Casajus (2018a) suggests a relaxation of symmetry, sign symmetry, that can replace symmetry in Young’s characterization of the Shapley value via efficiency, symmetry, and strong monotonicity or marginality (Young, 1985). In view of the former results, the question naturally arises whether the classes of weighted Shapley values can be characterized by efficiency, strong monotonicity or marginality, and weak sign symmetry or superweak sign symmetry.

7. Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 388390901.

References