

INTUITIVE SOLUTIONS IN GAME REPRESENTATIONS:
THE SHAPLEY VALUE REVISITED

By

Pradeep Dubey

March 2018

COWLES FOUNDATION DISCUSSION PAPER NO. 2123



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

<http://cowles.yale.edu/>

Intuitive Solutions in Game Representations: The Shapley Value Revisited*

Pradeep Dubey[†]

3 March 2018

Abstract

We show that *any* transferable utility game can be represented by an assignment of facilities to the players, in which it is intuitively obvious how to allocate the total cost of the facilities. The intuitive solution in the representation turns out to be the Shapley value of the game, and thus serves as an alternative justification of the value.

Key Words: TU game, characteristic function, Shapley value, assignment, representation.

JEL Classification: C71, C72, D61, D63, D70, D79

1 Introduction

We show that *any* characteristic function v on player set N (also known as a transferable-utility, or TU, game on N) can be “represented” by an assignment of

- (i) facilities that each player needs to use;
- (ii) cost/benefit of each facility.

Moreover, in this representation, all those who use any given facility are identical there¹ (in that the same facility suffices for any number of its users, with no alteration

*It is a pleasure to thank John Geanakoplos, Ori Haimanko, Abraham Neyman and Shyam Sunder for helpful comments. (In particular, Shyam helped me to come up with the title.)

[†]Stony Brook Center for Game Theory; and Cowles Foundation for Research in Economics, Yale University

¹Since we are considering a representation, the more special its features, the better. However, our assumption that users are identical *vis-a-vis* any given facility, is without loss of generality. As was already shown in [1], and as will shortly become evident in our more general context, any scenario with non-identical users can be replicated with identical users by expanding the number of facilities.

in its cost/benefit).

For ease of exposition, we shall from now on refer only to costs, with the explicit understanding that when a cost $\gamma < 0$ (or, $\gamma > 0$) it corresponds to a cost (or, benefit) in the colloquial sense.

Then the problem arises: how “should” the total cost, of all facilities, be allocated amongst N ?

An intuitive solution comes almost unbidden to the mind: at any facility, allocate its cost equally among all its (identical) users.

It turns out that this solution coincides with the Shapley value of v and thus serves to define the value on the domain of *all* characteristic functions.²

Worthy of note is the fact that the intuitive solution emerges *just* from the assignment, without any consideration of coalitions in N , leave aside a characteristic function on N .

One may (if of a scholastic turn of mind) explicate two principles behind the intuition, both so compelling as to require little justification. The first embodies a form of *decentralization*: the cost allocation at any facility does not depend on what is happening at other facilities (i.e., who goes to them, what their costs are, even how many of them there are, etc.; it is “independent of irrelevant information” regarding such externalities). The upshot is that, at any facility, those who do not use it may simply be ignored (indeed, they are literally non-existent from the standpoint of the facility, since they do not even come to it).

The second principle embodies a form of *non-discrimination*: at any facility, users who are identical must be treated in the same manner, i.e., they must be allocated the same costs.

2 The Assignment

Let N denote a finite set of *users* and K a finite set of *facilities*. Each user $n \in N$ is assigned a nonempty subset $\psi(n) \subseteq K$ of facilities that he must use; and each facility $k \in K$ is assigned a *cost* $\gamma(k) \in \mathcal{R}$ that accrues to its users. (Here \mathcal{R} denotes the reals; and the fact that $\gamma(k)$ depends only on k , not on the person who is using k , is tantamount to the assumption that all the users of k are identical at k .)

The pair (ψ, γ) will be called an *assignment*. In what follows, N will be fixed, while K, ψ, γ are allowed to vary in order to generate different assignments. (One may view (ψ, γ) as a bipartite graph, with disjoint sets N, K ; an edge (n, k) iff $k \in \psi(n)$; and cost $\gamma(k)$ at each node $k \in K$.)

²It also serves to explain the Shapley value to those who are not of a mathematical turn of mind.

Given the assignment (ψ, γ) , the natural question arises: how “should” the total cost $\sum_{k \in K} \gamma(k)$ be allocated to N ? The *intuitive solution*, as was said, stipulates that at any facility k , its cost $\gamma(k)$ be equally divided among all its identical users. Indeed, this solution was invariably given as being obvious, by a random assortment of lay people that the author posed the question to — high school students, fashion designers, actors, businessmen, politicians, *et al.* — all of them perfectly innocent of game theory. Yet their solution was the “Shapley value” of the “characteristic function” induced by (ψ, γ) .

Let us recall the notion of a characteristic function. Denote by $\mathcal{C} = \{T : T \subseteq N, T \neq \emptyset\}$ the set of *coalitions* in N . Let G be the Euclidean space whose axes are indexed by the elements of \mathcal{C} . A *characteristic function* v is vector in G (equivalently, a function $v : \mathcal{C} \rightarrow \mathcal{R}$), with the interpretation that its component $v(T)$ represents the payoff that coalition T can *guarantee* to its members. If v satisfies the condition $v(S \cup T) \geq v(S) + v(T)$ whenever $S \cap T = \emptyset$, it is said to be *superadditive*. Our analysis holds for every $v \in G$, superadditive or not, though it is conventional to assume superadditivity.

We shall show below that *any* $v \in G$ is induced by an assignment (ψ, γ) of the kind we have described; in our parlance, v is *represented* by the assignment (ψ, γ) . Assuming this for now, let us invoke the long forgotten “Axiom I” from Shapley’s 1951 working paper [3] (where “value” refers to the allocation of the total cost among the players). It is important to first stress that the domain, on which the value is sought to be defined in [3], does *not* consist of characteristic functions, but of what we call “scenarios” below (and Shapley called³ “games” in [3]).⁴

Axiom I: The (Shapley) value (in any scenario) depends only on the characteristic function v (that the scenario induces).

In light of this axiom, and our result that *all* characteristic functions on N can be represented by assignments, the intuitive solution (endorsed by our game-theoretic laity) determines the Shapley value on all scenarios (and of course — without the need for this axiom — on characteristic functions). See Section 7 for more details.

Axiom I seems to have sunk into the subconscious of the subject, and indeed the value is routinely taken to depend on v *by definition* (as is the case even in a subsequent version [4] of [3]). However the axiom does have significant content, which

³However, since characteristic functions have subsequently come to be commonly called “transferable utility (TU) games”, we use the term “scenarios” to prevent confusion.

⁴Scenarios constitute “representations” of the characteristic function they induce, even though they remained in the wings in [3] and only the characteristic function needed to be brought onto the stage.

is what presumably had led Shapley to state it in the first place. It calls attention to the fact that that, in different scenarios $\Gamma^1, \Gamma^2, \dots$, the value does not depend on the finer economic or engineering structure of the scenario, but only on its emergent “financial structure”, i.e., on its characteristic function v . (To put it differently: for purposes of cost allocation, what matters are the costs incurred by coalitions, never mind *how* they incurred it.) This opens up the possibility of *reversing the gaze*: rather than go from a scenario Γ to the characteristic function v that Γ induces, and then figure out the value to players in v , one could instead start with v and conjure up a special scenario Γ^* which represents v and in which the value is intuitively obvious.

3 The Characteristic Function Induced by an Assignment

Recall $\mathcal{C} = \{T : T \subseteq N, T \neq \emptyset\}$ is the set of coalitions in N . For any coalition $S \in \mathcal{C}$, let $\psi(S) = \cup_{n \in S} \psi(n)$ denote the set of facilities assigned to users in S , and let $v(S)$ be the cost of $\psi(S)$, i.e.,

$$v(S) = \sum_{k \in \psi(S)} \gamma(k)$$

This defines the *characteristic function* (or, *TU-game*) v induced by (ψ, γ) .

4 The Representation of a Characteristic Function by an Assignment

Definition 1 We say that the assignment (ψ, γ) is a representation of the characteristic function v if v is induced by (ψ, γ) .

Definition 2 For any assignment (ψ, γ) , let $|\psi^{-1}(k)|$ denote the cardinality of the set $\psi^{-1}(k) = \{n \in N : k \in \psi(n)\}$. The intuitive solution to (ψ, γ) is given by $\tau(\psi, \gamma) = (\tau_n(\psi, \gamma))_{n \in N}$ where

$$\tau_n(\psi, \gamma) = \sum_{k \in \psi(n)} \frac{\gamma(k)}{|\psi^{-1}(k)|}$$

Theorem 3 Let $v \in G$ be any characteristic function. There exists a representation of v by an assignment (ψ, γ) . Moreover, for any assignment (ψ', γ') that is a representation of v , the intuitive solution $\tau(\psi', \gamma')$ is the Shapley value of v .

Proof. Denote $\mathcal{C}^* = \{S \subseteq N : S \neq N\}$ (notice $\emptyset \in \mathcal{C}$), and for any $S \in \mathcal{C}^*$ define $w_S \in G$ by

$$w_S(T) = 0 \text{ if } T \subseteq S, \text{ and } w_S(T) = 1 \text{ otherwise}$$

We claim that the $2^n - 1$ vectors $\mathcal{W} = \{w_S : S \in \mathcal{C}^*\}$ form a basis for G . To see this, first define $u_S \in G$ for any $S \in \mathcal{C}$, by

$$u_S(T) = 1 \text{ if } T = S, \text{ and } u_S(T) = 0 \text{ otherwise}$$

i.e., $\{u_S : S \in \mathcal{C}\}$ is the standard basis of G consisting of the unit vectors. It suffices to show that each $u_S \in \text{Span } \mathcal{W}$. We do so by induction on the cardinality s of S . First let $s = 1$. Then $S = \{i\}$ for some $i \in N$, and clearly

$$u_{\{i\}} = w_\emptyset - w_{\{i\}} \in \text{Span } \mathcal{W}$$

Next assume $u_S \in \text{Span } \mathcal{W}$ for $s = 1, \dots, k$. Let S have cardinality $k + 1$ and let S_1, \dots, S_l denote all the *strict* subsets of S . Consider

$$w^* = w_\emptyset - w_S$$

In w^* observe that S and its strict subsets get 1, and all other coalitions get 0. Therefore

$$u_S = w^* - u_{S_1} - \dots - u_{S_l}$$

Note $w^* \in \text{Span } \mathcal{W}$ by the second-last display, and $u_{S_j} \in \text{Span } \mathcal{W}$ for $j = 1, \dots, l$ by the inductive assumption (since the cardinality of each S_j is k or less). Hence the last display implies that $u_S \in \text{Span } \mathcal{W}$. This establishes the claim.

Now consider any $v \in G$. By the claim, there are unique scalars c_S such that

$$v = \sum_{S \in \mathcal{C}^*} c_S w_S$$

Let the set K of facilities correspond to \mathcal{C}^* and define the assignment (ψ, γ) as follows⁵:

$$\psi(n) = \{S \in \mathcal{C}^* : n \notin S\}, \text{ for } n \in N;$$

and

$$\gamma(S) = c_S, \text{ for } S \in \mathcal{C}^*.$$

⁵If $c_S = 0$, one could of course drop facility S .

It is readily verified that (ψ, γ) represents v . Indeed, by the linear expression for v (third last display), we have

$$v(T) = \sum_{S \in \mathcal{C}^*} c_S w_S(T)$$

for any $T \in \mathcal{C}$. But $w_S(T) = 1$ if, and only if, $T \not\subseteq S$; therefore

$$v(T) = \sum_{S \in \mathcal{C}^*, T \not\subseteq S} c_S$$

Since the users in T are linked by ψ to precisely those facilities in $S \in \mathcal{C}^*$ such that $T \not\subseteq S$ (see the fourth last display), i.e., $\psi(T) = \{S \in \mathcal{C}^* : T \not\subseteq S\}$, it follows that (ψ, γ) represents v .

Next let $\varphi : G \rightarrow \mathcal{R}^N$ denote the Shapley value, and let (ψ', γ') be any representation of $v \in G$ with K' as its underlying set of facilities. For any $k \in K'$, define the TU game $v_k \in G$ by

$$v_k(T) = \gamma'(k) \text{ if } k \in \psi'(T); \text{ and } v_k(T) = 0 \text{ otherwise}$$

where (recall) $\psi'(T) = \cup_{n \in T} \psi'(n)$. Then, from the fact that (ψ', γ') represents v , it is easy to check that

$$v = \sum_{k \in K'} v_k$$

Thus we have $\varphi(v) = \sum_{k \in K'} \varphi(v_k)$ by the additivity of φ . But, in the TU game v_k , all the players in $T_k = \{n \in N : k \notin \psi'(n)\}$ are dummies and all the players in $N \setminus T_k$ are symmetric. Hence $\varphi(v_k)$ awards 0 to players in T_k , and divides the total $v_k(N) = \gamma'(k)$ equally among the players in $N \setminus T_k$ (using the dummy, symmetry and efficiency properties of φ). This implies that $\varphi(v)$ is the intuitive solution $\tau(\psi', \gamma')$.

■

5 A Narrative for the Assignment

We present here a narrative that elaborates on how a characteristic function v is induced by an assignment (ψ, γ) in accordance with our definition. For simplicity, let us focus on the case where the induced v is superadditive (this is an implied restriction on pairs (ψ, γ) .)

Assume first that each $n \in N$ *must* use facilities in $\psi(n)$ – no more and no less – otherwise a huge penalty (say, $-\infty$) is levied on n . Next assume that the only way for a user to obtain access to facilities is for him to be part of a coalition which registers itself with a central authority (coalitions can be singletons). Once a coalition S registers, the authority allocates to S an “industrial park”, indexed by S , at which only the members of S are allowed to enter and at which precisely all the facilities in $\psi(S)$ are available (along with their attendant costs given by γ). Now a little reflection reveals that the best that S can *guarantee* to itself, is to go as a block to the authority, since there is no incentive for S to split into smaller subsets (on account of superadditivity), and since *one* of the many courses of actions (rational or not) available to those in $N \setminus S$ is to go off on their own. (By going as a block, S incurs the cost $v(S) = \sum_{k \in \psi(S)} \gamma(k)$; while by going off, $N \setminus S$ can ensure that S does no better.)

When v is not superadditive, one might need to strengthen the narrative, and further assume that once players enter any park S they can never leave it, or interact in any way with others outside of the park. (We skip the details.)

The point is not whether the above narrative is realistic or interesting. It may not even be happening on earth but on some other (to our prejudiced minds) wierd planet. All that matters is that the narrative be logically coherent, with no internal contradictions, and that the numbers $v(S)$ be well-defined for every $S \in \mathcal{C}$, compatible with the interpretation

$v(S) =$ maximum payoff that S can guarantee itself, no matter what players in $N \setminus S$ do

We admit any such narrative in (ψ, γ) as a scenario in the domain to which Axiom 1 applies.

6 Equivalent Assignments

Our analysis implies that the intuitive solution (which may be summed up in the simple phrase “equal division, facility-by-facility”) is invariant across different assignments $(\psi, \gamma), (\psi', \gamma'), (\psi'', \gamma''), \dots$, provided that these assignments are representations of the same v . This may be of some independent interest in network analysis.⁶

⁶I owe this observation to John Geanakoplos.

7 An Axiomatic Framework

Let \mathfrak{S} be an abstractly given set of “scenarios” on player-set N . Here by a scenario we mean any interaction among the players, in which it is clear how to define the amounts $v(S)$ that coalitions $S \in \mathcal{C}$ can guarantee to their members (see Section 5). For our purposes, it is not necessary to be more specific.

Further, let \mathfrak{A} denote the set of *all* assignment scenarios on N as defined earlier; and recall that $\tau : \mathfrak{A} \rightarrow \mathcal{R}^N$ denotes the intuitive solution on \mathfrak{A} .

Finally let $\theta : \mathfrak{S} \rightarrow \mathcal{R}^N$ be a payoff map to the players as in [3].

Axiom II: $\mathfrak{A} \subset \mathfrak{S}$ and θ coincides with τ on \mathfrak{A} .

In other words, the domain of θ is rich enough to include every assignment, and on each assignment it prescribes the intuitive solution. Our theorem may then be restated: *the unique map θ which satisfies Axioms I and II is the Shapley value.*

As was pointed out, Axiom II can be split into two parts, requiring θ to satisfy decentralization and non-discrimination. Also note that Axiom I is only needed if \mathfrak{S} extends beyond \mathfrak{A} .⁷

8 Related Literature

By way of a concrete instance of the assignment (ψ, γ) , think of N as a set of identical planes, and of $\psi(n)$ as the set of cities that plane n must fly to.⁸ If $\gamma(k) < 0$, it represents the cost of building the runway at city k (a runway, once built, accomodates any number of planes since they are of identical make); and if $\gamma(k) > 0$, one may think of it as a “subsidy” given by city k (collectively to the planes that visit its remote location). (If $\gamma(k) = 0$, city k can be dropped from the picture.) These costs (benefits) may vary across cities on account of different costs of labor, land, location, etc.

This picture is based on [2], which focused attention on a single airport used by different-sized planes; and on [1] which observed that the model of [2] can be recast with identical-sized planes that have different flight paths among multiple airports. However the analysis in [1] (as in [2]) was restricted to the case of costs $\gamma(k) < 0$, which severely limited the class of characteristic functions that could be generated. The fact was missed in [1] that, by letting $\gamma(k)$ take on arbitrary values in \mathcal{R} , one can generate all possible characteristic functions.

⁷This section was prompted by a query from Ori Haimanko.

⁸The daily flight path of each plane is a loop which traverses the cities in $\psi(n)$ (in any order), taking off and landing at each city in $\psi(n)$ exactly once.

References

- [1] Dubey, P. (1982). The Shapley Value as Aircraft Landing Fees – Revisited, *Management Science*, Vol. 28, Issue 8, pp 869-874.
- [2] Littlechild, S.C. and G.F.Thompson (1977). Aircraft Landing Fees: A Game Theory Approach. *The Bell Journal of Economics*, Vol. 8, No.1, pp 186-204.
- [3] Shapley, L.S. (1951). Notes on the n-Person Game – II: The Value of an n-Person Game, The RAND Corporation, *The RAND Corporation*, Research Memorandum 670.
- [4] Shapley, L.S. (1953). A Value for n-Person Games, in *Contribution to the Theory of Games*, Vol. II, ed. H.W.Kuhn & A.W.Tucker, Annals of Math. Studies 28, Princeton University Press, NJ, pp 307-17