A BARGAINING PROCEDURE

LEADING TO THE SHAPLEY VALUE

Morton Davis

Econometric Research Program
Research Memorandum No. 61
December 3, 1963

The research described in this paper was supported by The Carnegie Corporation of New York.

Princeton University
Econometric Research Program
92-A Nassau Street
Princeton, N.J.

Abstract

In a paper by Shapley, "A Value for n-person games" there is assigned to each player the expected value or worth of an n-person game given in characteristic function form. This is done both by an axiomatic approach and from a bargaining model. A more complex bargaining model is constructed here which also leads to the Shapley Value, and the limitations of this model are considered.

§1. Introduction

In [1], L. Shapley derives a formula for the value of a game (given in characteristic function form) to a participant. The derivation is made first (and primarily) from three axioms and the result is then interpreted as follows: If an imputation is formed (i.e., a coalition of all the players) one player at a time, and at each stage of the formation the additional player gets the marginal value he adds (i.e., the difference between the coalition value after he enters and the coalition value just before he enters) and if we assume the n-person coalition is as likely to form in one way as another, the Shapley Value for a player may be defined as the expected value of a player's gain.

There is some question, however, whether one coalitional formation is as likely to occur as any other and whether, during such a formation, the marginal player would get all the marginal wealth. These two questions are related, of course, since the likelihood of a coalition forming is a function of what payoffs are made.

{I appreciate that in the original paper by Shapley [1] the bargaining model was only incidental to the derivation of the value from the axioms and the problems raised here were not really considered seriously}.

This paper is an attempt to deal with both of the above problems (an attempt which has achieved only partial success) by giving a bargaining model which is hopefully more plausible, and which in another way indicates that the Shapley Value is a reasonable a priori expectation.

§2. Definitions and Motivation

The game we will deal with consists of a universe of players: N, and a function, V, defined on the subsets of N, to the non-negative, real numbers and called the characteristic function. The characteristic function is assumed to be superadditive, i.e.,

$V(S) \geq V(S \cap T) + V(S-T)$ for all $S, T \subseteq N$ and

$V(\Phi) = 0$. We denote by

$N \equiv$ the set of players $\{1,2,\ldots n\}$

and $n \equiv$ the number of players in the game. In general we denote by capital letters the subsets of N and by small letters the number of elements of the subset, e.g., s = the number of elements in $S \subseteq N$.

Following L. Shapley [1] we will also attempt to define a vector valued function $\Phi[V]$ where the argument of the function is the characteristic function of the game and the range is an n-dimensional vector which has as its i^{th} component the "value" to the i^{th} player.

Our derivation of the value, however, will be from the bargaining model. Heuristically, the model we have in mind is as follows.

We assume that ultimately an imputation is formed, i.e., the players are all in one coalition. The universal coalition is not formed immediately however, but is formed by a player (set of players) joining another in a sequence of discrete steps. When a player first joins another, there is a certain non-negative gain achieved by superadditivity, i.e., the coalition value of the set consisting of two players (or two subsets) has a characteristic function value not less than the characteristic function values of the two player (or subsets) added. It is assumed that this extra wealth at this stage is divided among the two players in a way which reflects their competitive strengths. We choose to measure the competitive strength of a player i vis-a-vis a player j by asking what the value to the player i would be of the game (modified in the obvious way) with j excluded, asking the same question with the roles of players i and j reversed and taking the difference. There is nothing compelling about this criterion but it does not seem unreasonable. We shall refer to a player's share in an intermediate coalition as his fair share. Once this intermediate coalition is made, the bargaining procedure continues but with the following changes: the two players act as one; they subsequently share any further gains equally and the characteristic function value of each coalition to which the two belong is reduced by the amount of the characteristic value of this two person coalition: the amount they have withdrawn from the game. By reducing the computation of the value to a player of an n-person game to that of the computation of the value to a player of an n-1 person game we may using mathematical induction find the value to a player of any game once we postulate what the value to a player of a two person game is: in normalized form where each player alone gets nothing, the value of the game to each player is half the characteristic function value for the two person Coalition.

The new value we get (which turns out to be the Shapley Value) must also be considered as an a priori value made under the assumption that each path of forming the ultimate imputation is as likely as any other. The reasonableness of this assumption will be considered later.

§3. Derivation of the Main Theorem

We give our Axioms for the Bargaining model together with an intuitive explanation below.

AXIOM I In the two Person Game consisting of players i and j the fair share of player i, j, is $1/2\,[V(i\ j) + V(i) - V(j)]$, $1/2[V(i\ j) + V(j) - V(i)]$, respectively.

AXIOM II Assume we have for any (N-1) person game in characteristic function form a share defined for each of the players. We will describe the calculation of the fair share of a player i in an n-person game (and give a heuristic justification in terms of the dynamics of the bargaining model, simultaneously.

Case 1 - Player i joins player j to form an initial two person coalition and $V(i\ j)$ is distributed among them in a way that reflects their relative bargaining strengths. Formally, we define

y_i^{ij}, y_j^{ij} so that (A) and (B) below are satisfied.

$$\text{(A)} \quad y_i^{ij} + y_j^{ij} = V(i\ j)$$

$$\text{(B)} \quad y_i^{ij} - y_j^{ij} = x_i^j - x_j^i$$

where y_i^{ij}, y_j^{ij} are the respective shares of players i, j in the initial two person coalition $\{i,j\}$ and x_i^j is the share of player i in the N-1 person game consisting of $N - \{j\}$ and with characteristic function identical to the original characteristic function but of course defined only for sets not containing player j. (The existence of such a fair share is ensured by the inductive hypothesis). x_j^i is defined analagously.

Define Z_{ij} to be the fair share of the coalition (i,j) in the original game where the coalition (i,j) acts as a single player and takes the $V(ij)$ for itself in advance.

Formally, we define a new game consisting of the n-1 players $N-\{i,j\}\cup\{\alpha\}$ with a new characteristic function W defined over the subsets of $N-\{i,j\}\cup\{\alpha\}$

where $W(A) = V(A)$ if $\alpha \notin A$

and $W(A) = V(A) - V(i,j)$ if $\alpha \in A$;

then Z_{ij} = the fair share of the player α, i.e.,

$$\text{(C)} \quad Z_{ij} = \sum_{\substack{i,j \in S \\ S \subseteq N}} \gamma_{n-1}(s-1)\,[V(S)-V(S-\{i,j\})]$$

which is in terms of the theorem given below.

Finally, x_i^{ij} : player i's fair share if the pair (i,j) is formed initially, is defined to be

$$\text{(D)} \quad x_i^{ij} = y_i^{ij} + Z_{ij}/2$$

<u>Case 2</u> Player j and k join initially (both distinct from i). They act as a single player in the same way that i and j did in the earlier case. Formally, we define x_i^{jk} to be the fair share of player i in the n-1 person game

$N - \{j,k\} \cup \{\alpha\}$ with characteristic function W defined below:

$$W(A) = V(A) \quad \alpha \notin A$$

$$W(A) = V(A) - V(jk) \quad \alpha \in A$$

We now are in a position to define the fair share of player i, x_i, in the n-person game with characteristic function V. x_i is the expected value of player i if each two person coalition is formed with equal probability. Formally,

$$\text{(E)} \quad x_i = \frac{\sum_{j \neq i} x_i^{ij} + \sum_{j<k} x_i^{jk}}{\frac{n(n-1)}{2}} \qquad i,j,k \text{ distinct}$$

<u>Comments:</u> There are two comments that should be made at this point. First, the y_i^{ij} may conceivably be negative which seems to violate individual rationality. This might occur if a player joined another much stronger than himself and payed a premium for the privilege, i.e., the increased bargaining strength of the new alliance. This does not seem unreasonable although in terms of this model (which is of a stochastic nature) he may well receive less than zero at the end.

Secondly, the choice of equation (B) to reflect the bargaining strength of player i vis-a-vis player j is somewhat arbitrary (see references [2] and [3] where other criteria were used) but I feel not unreasonable.

THEOREM The fair share of player i, x_i, is the Shapley Value for player i, i.e.,

$$x_i = \Phi_i(V) = \sum_{S \text{ a subset of } N} \gamma_n(s)\,[V(S) - V(S-\{i\})] \quad \text{where}$$

s is the number of players in coalition S and

$$\gamma_n(s) = \frac{(s-1)!\;(n-s)!}{n!}$$

Proof: The proof is by induction. It is clearly true for $n = 2$ by axiom I.

Let us assume that the theorem is true for every $(n-1)$-person game; we show below it is true for every n-person game.

We note first, using the inductive hypothesis that by (A), (B)

$$y_i^{(ij)} = \frac{1}{2}\,[V(i,j) + x_i^{\,j} - x_j^{\,i}]\,, \quad i \neq j$$

$$x_i^{\,j} = \sum_{S \text{ a subset of } N-\{j\}} \gamma_{n-1}(s)\,[V(S) - V(S - \{i\})]$$

$$x_j^{\,i} = \sum_{S \text{ a subset of } N-\{i\}} \gamma_{n-1}(s)\,[V(S) - V(S - \{j\})]$$

Also using the inductive Hypothesis again and (C)

$$Z_{ij} = \sum_{i,j \in S} \gamma_{n-1}(s-1)\,[V(S) - V(S - \{i,j\})]\ .$$

By (D)

$$x_i^{\,ij} = y_i^{\,ij} + \frac{Z_{ij}}{2} \quad \text{or}$$

(F)

$$x_i^{\,ij} = \frac{1}{2}\,\Big[\sum_{i,j \in S} \gamma_{n-1}(s-1)\,V(S) + \sum_{\substack{i \in S \\ j \notin S}} \gamma_{n-1}(s)\,V(S) - \sum_{\substack{j \in S \\ i \notin S}} \gamma_{n-1}(s)\,V(S) - \sum_{i,j \notin S} \gamma_{n-1}(s+1)\,V(S)\Big]$$

and summing over $j \in N$, $j \neq i$ we have,

$$\sum_{j\neq i} X_i^{ij} = \frac{1}{2} [\Sigma \{(s-1)\, \gamma_{n-1}(s-1)\, V(S) + (n-s)\, \gamma_{n-1}(s)\, V(S)\}$$

$$- \sum_{i\notin S} \{s\gamma_{n-1}(s)\, V(S) + (n-s-1)\, \gamma_{n-1}(s+1)\, V(S)\}]$$

Noting that

$$(s-1)\, \gamma_{n-1}(s-1) + (n-s)\, \gamma_{n-1}(s)$$

$$= \frac{(s-1).(s-2)!.\ (n-s)!}{(n-1)!} + \frac{(n-s)\ (s-1)!\ (n-s-1)!}{(n-1)!}$$

$$= \begin{cases} 2n\, \gamma_n(s) & \text{if } S \neq N \\ 1 & \text{if } S = N \end{cases}$$

and also,

$$s\, \gamma_{n-1}(s) + (n-s-1)\, \gamma_{n-1}(s+1)$$

$$= \frac{s\ (n-s-1)!\ (s-1)!}{(n-1)!} + \frac{(n-s-1)\ (n-s-2)!\ s!}{(n-1)!}$$

$$= \begin{cases} 2n\, \gamma_n(s+1) & \text{if } s<n-1 \\ 1 & \text{if } s=n-1 \end{cases}$$

and we have

$$\sum_{j\neq 1} X_i^{ij} = n \Big[\sum_{\substack{i\in S \\ S\neq N}} \gamma_n(s)\, V(S) - \sum_{\substack{i\notin S \\ S\neq N-\{i\}}} \gamma_n(s+1)\, V(S)\Big] + \frac{1}{2} [V(N) - V(N-\{i\})] .$$

(G) $$= n \Big[\sum_{\substack{i\in S \\ S\neq N}} \gamma_n(s)\, [V(S) - V(S-\{i\})]\Big] + \frac{1}{2} [V(N) - V(N-\{i\})]$$

On the other hand if (j,k) form the first two person coalition then

$$x_i^{jk} = \sum_{\substack{i \in S \\ j,k \notin S}} \gamma_{n-1}(s) [V(S) - V(S-\{i\})] + \sum_{i,j,k \in S} \gamma_{n-1}(s-1) [V(S)-V(S-\{i\})]$$

(where i, j, k are distinct)

$$= \sum_{i,j,k \in S} \gamma_{n-1}(s-1) V(S) + \sum_{\substack{i \in S \\ j,k \notin S}} \gamma_{n-1}(s) V(S)$$

$$- \sum_{i,j,k \notin S} \gamma_{n-1}(s+1) V(S) - \sum_{\substack{j,k \in S \\ i \notin S}} \gamma_{n-1}(s) V(S) ;$$

Summing over j, k, we have

$$\sum_{\substack{j<k \\ i,j,k \text{ distinct}}} x_i^{jk} = \sum_{\substack{i \in S \\ S \neq N}} [\gamma_{n-1}(s-1) \frac{(s-1)(s-2)}{2} + \gamma_{n-1}(s) \frac{(n-s)(n-s-1)}{2}] V(S)$$

$$- \sum_{\substack{i \notin S \\ S \neq N-\{i\}}} [\gamma_{n-1}(s+1) \frac{(n-s-1)(n-s-2)}{2} + \gamma_{n-1}(s) \frac{s(s-1)}{2}] V(S) + \frac{n-2}{2} [V(N)-V(N-\{i\})]$$

$$= \sum_{\substack{i \in S \\ S \neq N}} [\frac{(s-1)(s-2)}{2} \frac{(n-s)!(s-2)!}{(n-1)!} + \frac{(n-s)(n-s-1)}{2} \frac{(n-s-1)!\ (s-1)!}{(n-1)!}] V(S)$$

$$- \sum_{\substack{i \notin S \\ S \neq N-\{i\}}} [\frac{(n-s-1)(n-s-2)(n-s-2)!s!}{2\ (n-1)!} + \frac{s(s-1)(n-s-1)!(s-1)!}{2\ (n-1)!}] V(S) + \frac{n-2}{2} [V(N)-V(N-\{i\})]$$

$$= \sum_{\substack{i \in S \\ S \neq N}} \frac{n(n-3)}{2} \gamma_n(s)V(S) - \sum_{\substack{i \notin S \\ S \neq N-\{i\}}} \frac{n(n-3)}{2} \gamma_n(s+1) V(S) + \frac{n-2}{2} [V(N-\{i\})] ,$$ So,

$$\text{(H)} \quad \sum_{\substack{j<k \\ i,j,k \text{ distinct}}} x_i^{jk} = \frac{n(n-3)}{2} \sum_{i \in S} \gamma_n(s) [V(S) - V(S-\{i\})] + \frac{n-2}{2} [V(N) - V(N-\{i\})]$$

If we substitute (G) and (H) into (E) we have

$$x_i = \frac{\sum_{j \neq i} x_i^{ij} + \sum_{\substack{j<k \\ i,j,k \text{ distinct}}} x_i^{jk}}{\frac{n(n-1)}{2}}$$

$$= \frac{\left(n + \frac{n(n-3)}{2} \right) \sum_{\substack{i \in S \\ S \neq N}} \gamma_n(s) \left[V(S)-V(S-\{i\}) \right] + \left(\frac{1}{2} + \frac{n-2}{2} \right) \left[V(N)-V(N-\{i\}) \right]}{\left(\frac{n(n-1)}{2} \right)}$$

$$= \sum_{i \in S} \gamma_n(s) \left[V(S)-V(S-\{i\}) \right] \quad \left(\text{since } \frac{1}{n} = \gamma_n(n) \right)$$

, which was to be shown.

§4 Conclusion

A detailed computation and commentary on the three person case will be profitable despite its simplicity. We denote the three players by i, j, k.

Let us assume i and j first combine;

Then $y_i^{ij} + y_j^{ij} = V(ij)$

$$y_i^{ij} - y_j^{ij} = x_i^{j} - x_j^{i} = \frac{V(ik) - V(jk)}{2}$$

hence

$$y_i^{ij} = \frac{2V(ij) + V(ik) - V(jk)}{4}$$

also, $Z_{ij} = \dfrac{V(ijk) - V(ij)}{2}$

so $x_i^{ij} = \dfrac{V(ijk) + V(ij) + V(ik) - V(jk)}{4}$

and $x_k^{ij} = \dfrac{V(ijk) - V(ij)}{2}$;

These calculations are sufficient to tell the whole story since there is only one intermediate coalition to be formed and a player is either in or out.

B2

The first thing we note is that x_i^{ij} is symmetric in j and k, i.e., $x_i^{ij} = x_i^{ik}$. It appears, then, that the final payoff to a player depends solely upon whether he enters into an initial two person coalition or not. This enables us to consider the "value" assigned to each player in this game as the weighted average of two other "values": the "value" to a player of the game should he be in the initial two person coalition multiplied by two thirds (since he will be in it presumably two thirds of the time) plus the "value" to the player of the game if he is not in this initial two person coalition multiplied by one third. The former in-coalition "value" will be determined by the amount the player demands in his initial two person coalition which in turn will be determined by competitive equilibrium in the manner of [4] Pg. 228, [2], and [3]. The out-coalition player will have, to a certain extent, his value foisted upon him. It might be hoped, then, that in the general n-person game we could have each player i demand from each other player j such a part of $V(ij)$ that this share plus his subsequent expectation would be invariant with respect to j , just as in the three person case it mattered not what coalition a player entered but merely whether he entered or did not enter an initial coalition. We will attempt to construct such a model. In particular we will assume there is a value $W(i)$ to which player i may aspire, i.e. the share of $V(ij)$, α_i^{ij} player i demands (perhaps negative) from player j in joining the initial coalition $(i\ j)$ (acting subsequently as a single player) will be the same regardless of which j he joins. Formally,

Define: α_i^{ij} = The share of $V(ij)$ i gets in the initial formation of the two person coalition (i,j). K_{ij} = the expected value of the gain of the coalition (ij) acting as a single player in a new modified (n-1) person game with the characteristic function defined as earlier. (The value of K_{ij} may be deduced by induction once we assume the division of the two person game, exactly as was done earlier). β_i^{ij} = the fractional share of K_{ij} to which player i is entitled. We may then write

$$\text{(i)}\quad W(i) = \alpha_i^{ij} + \beta_i^{ij} K_{ij} \quad \text{for } i = j,\ 1 \leq i,j, \leq n \ ;$$

$$\text{(ii)}\quad \alpha_i^{ij} + \alpha_j^{ij} = \beta_i^{ij} + \beta_j^{ij} = 1$$

For the three person case we have

$$\alpha_1^{12} + \beta_1^{12} K_{12} = W(1) = \alpha_1^{13} + \beta_1^{13} K_{13}$$

$$\alpha_2^{23} + \beta_2^{23} K_{23} = W(2) = \alpha_2^{21} + \beta_2^{21} K_{12}$$

$$\alpha_3^{13} + \beta_3^{13} K_{13} = W(3) = \alpha_3^{23} + \beta_3^{23} K_{23}$$

where (iii) $K_{ij} = \frac{1}{2} [\, V(N) - V(ij) \,]$

by our earlier assumtion about the two person case.

Solving (i) using (ii) and (iii)

we obtain $W(1) = \alpha_1^{12} + \beta_1^{12} K_{12}$

$$= \frac{1}{2} [\, K_{12} + K_{13} - K_{23} + V_{12} + V_{13} - V_{23} \,] = \frac{1}{4} [\, V(N) + V(12) + V(13) - V(23) \,]$$

and similarly,

$$W(2) = \frac{1}{4} [\, V(N) + V(12) + V(23) - V(13) \,]$$

$$W(3) = \frac{1}{4} [\, V(N) + V(13) + V(23) - V(12) \,]$$

It is interesting that in this formulation of the problem we arrive at the same solution that we did in the earlier case where an apparently more specific procedure was prescribed. Unfortunately, by examining the set of equations above in the case of a four person or larger game it is easy to check that they will be overdetermined, i.e., the number of equations will be greater than the number of unknowns and thence will in general be inconsistent. Without going into the mathematical details the intuitive reasons for this will be indicated: each pair of players, if they form first, have a value which is actually a sum of two values: V(ij) plus what (ij) may make as a single player which the two must divide amongst themselves. This will yield equations of the type.

$$[\text{ value of } (i,j)\ ] - W(j) = [\text{value of } (i,k)\ ] - W(k)$$

where k,i,j are distinct. This yields $1+2+\ldots\ldots(n-1) = \frac{n(n-1)}{2}$ equations with only n unknowns: W(i) , i = 1 to n.
These two values are equal if and only if n=3 (or n=o which is meaningless in this context).

Hence, for the general n-person game using our bargaining model, it appears one may not demand that each player be indifferent about which coalition he enters.

It follows that there is an order preference for a player with respect to which players he would like to join to form a preliminary coalition and this in turn calls into question the assumption that any coalition is as likely to form as any other amd on that basis computing the "value" to a player of the game.

REFERENCES

1. L. S. Shapley, "A Value for n-person games", "Annals of Mathematics Study No. 28 (Princeton 1953) pp. 307-317.

2. R. J. Aumann and M. Maschler, "The Bargaining Set for Cooperative Games". To appear in Advances in Game Theory, M. Dresher, L. S. Shapley and A. W. Tucker, eds. ANNALS OF MATHEMATICS STUDIES, No. 52, Princeton University Press, Princeton, New Jersey.

3. M. Davis and M. Maschler, "The Kernel of a Cooperative Game". Econometric Research Memorandum No. 58, June 20, 1963.

4. J. Von Neumann and O. Morgenstern, "Theory of Games and Economic Behavior", Princeton 1944, 3rd ed. 1953.