Roth, A.E.,
"The Shapley Value as a von Neumann- Morgenstern Utility,"
Econometrica, 45, April 1977, 657-664.

Econometrica, Vol. 45, No. 3 (April, 1977)

THE SHAPLEY VALUE AS A VON NEUMANN-MORGENSTERN UTILITY

BY ALVIN E. ROTH[1]

The Shapley value is shown to be a von Neumann–Morgenstern utility function. The concept of *strategic risk* is introduced, and it is shown that the Shapley value of a game equals its utility if and only if the underlying preferences are neutral to both ordinary and strategic risk.

1. INTRODUCTION

THE DEVELOPMENT OF GAME THEORY has been closely associated with the axiomatic treatment of cardinal utilities, ever since both were introduced by von Neumann and Morgenstern. In *Theory of Games and Economic Behavior* [6] they give a set of axioms which assure that a cardinal utility function can be deduced from ordinal preferences. This cardinal utility is then used to define the characteristic function form of a game.

In 1953, in his classic paper "A Value for *n*-Person Games," [5], L. S. Shapley notes that this cardinal utility is defined only for "simple situations"—prizes, and lotteries over prizes—but not for games themselves. He says:

> "At the foundation of the theory of games is the assumption that the players of a game can evaluate, in their utility scales, every "prospect" that might arise as a result of a play. In attempting to apply the theory to any field, one would normally expect to be permitted to include, in the class of "prospects," the prospect of having to play a game. The possibility of evaluating games is therefore of critical importance. So long as the theory is unable to assign values to the games typically found in application, only relatively simple situations—where games do not depend on other games—will be susceptible to analysis and solution."

He proceeds to give three cardinal conditions which a value for games should satisfy, and to show that there is a unique function which satisfies these conditions.

In this paper, we develop a cardinal utility function for games, based on ordinal preferences. Our treatment shall differ from the elementary axiomatization of utilities from preferences in that games—the objects over which our preferences are defined—are themselves defined in terms of a cardinal utility.

Thus, it shall be necessary to insure that the utility function for games is compatible with the existing utility function which defines the games. We shall show that this requirement leads to a unique utility function for games, and that, when the underlying preference relation is neutral with respect to certain kinds of risk, the utility of a game is equal to the Shapley value. Our approach will be to develop a preference relation which permits us to compare positions in a game and in different games, by extending the utility function used to define the games.

[1] The author wishes to acknowledge some thoughtful comments by David Schmeidler.

2. UTILITY THEORY

We summarize here an axiomatization of utility presented in [**1**].

A set M is a *mixture set* if for any elements $a, b \in M$, and for any number $p \in [0, 1]$, we can associate another element of M denoted by $[pa; (1-p)b]$ called a *lottery* between a and b. (Henceforth the letters p and q will be reserved for elements of $[0, 1]$.) We assume that lotteries have the following properties for all $a, b \in M$:

$$[1a; 0b] = a, \quad [pa; (1-p)b] = [(1-p)b; pa], \quad \text{and}$$
$$[q[pa; (1-p)b]; (1-q)b] = [pqa; (1-pq)b]. \tag{2.1}$$

A *preference* on M is defined to be a binary relation $\succcurlyeq$ such that for any $a, b \in M$ either $a \succcurlyeq b$ or $b \succcurlyeq a$ must hold, and if $a \succcurlyeq b$ and $b \succcurlyeq c$ then $a \succcurlyeq c$. (We write $a \succ b$ if $a \succcurlyeq b$ and $b \not\succcurlyeq a$, and $a \sim b$ if $a \succcurlyeq b$ and $b \succcurlyeq a$.) A real valued function u defined on a mixture set M is a *utility function* for the preference $\succcurlyeq$ if it is order preserving (i.e., if $\forall a, b \in M$, $u(a) > u(b)$ if and only if $a \succ b$), and if

$$u([pa; (1-p)b]) = pu(a) + (1-p)u(b). \tag{2.2}$$

If $\succcurlyeq$ is a preference ordering on a mixture set M, then the following conditions insure that a utility function exists:

CONDITION 1: For any $a, b, c \in M$, the sets $\{p \mid [pa; (1-p)b] \succcurlyeq c\}$ and $\{p \mid c \succcurlyeq [pa; (1-p)b]\}$ are closed.

CONDITION 2: If $a, a' \in M$ and $a \sim a'$ then for any $b \in M$, $[\frac{1}{2}a; \frac{1}{2}b] \sim [\frac{1}{2}a'; \frac{1}{2}b]$.

The utility function is unique up to an affine transformation. For any element $x \in M$, the utility of x can be given by

$$u(x) = \frac{p_{ab}(x) - p_{ab}(r_0)}{p_{ab}(r_1) - p_{ab}(r_0)}$$

where a, b, r_1, and r_0 are elements of M such that $a \succcurlyeq x \succcurlyeq b$ and $a \succcurlyeq r_1 \succcurlyeq r_0 \succcurlyeq b$, and for any $y \in M$ such that $a \succcurlyeq y \succcurlyeq b$, $p_{ab}(y)$ is defined by

$$y \sim [p_{ab}(y)a; (1 - p_{ab}(y))b]. \tag{2.3}$$

It can be shown that the numbers $p_{ab}(\cdot)$ are well defined, and the function $u(\cdot)$ is independent of the choice of a and b. The fixed elements r_1 and r_0 determine the origin and scale of the utility function; note that $u(r_1) = 1$, and $u(r_0) = 0$.

3. THE SHAPLEY VALUE

We summarize here the axiomatization of the value presented in [**5**], giving first some necessary definitions.

Denote by N the universal set of *players* (or *positions*), and define a *game* to be any function v from the subsets of N to the real numbers such that for all sets $S, T \subseteq N$,

$$v(S) \geqslant v(S \cap T) + v(S - T) \quad \text{and} \quad v(\bar{0}) = 0,$$

where $\bar{0}$ denotes the empty set. The quantity $v(S)$ can be interpreted as the wealth obtainable by the coalition S, and the model implicitly assumes that wealth is transferable between players, and that utility is linear in wealth.

For convenience we shall assume that N is a finite set, and denote its cardinality by n. (Similarly, the cardinality of sets R, S, and T will be denoted r, s, and t.)

A *carrier* of v is any set $T \subseteq N$ such that for all $S \subseteq N$

$$v(S) = v(T \cap S).$$

The superset of any carrier is itself a carrier. If π is a permutation of N (i.e., if π is a one to one mapping from N to itself) then, for all sets $S \subseteq N$, we denote the image of S under π by πS, and define the game πv by

$$\pi v(\pi S) = v(S).$$

By the *value* of a game v we mean a vector valued function $(\phi_1(v), \phi_2(v), \ldots, \phi_n(v))$ which associates a real number $\phi_i(v)$ with each position $i \in N$, and which obeys the following conditions.

CONDITION 3: For each permutation π, $\phi_{\pi i}(\pi v) = \phi_i(v)$.

CONDITION 4: For each carrier T of v, $\Sigma_{i \in T} \phi_i(v) = v(T)$.

CONDITION 5: For any games v and w, $\phi(v + w) = \phi(v) + \phi(w)$.

Note that Condition 5 implies that if $(v - w)$, v, and w are all games, then $\phi(v - w) = \phi(v) - \phi(w)$.

Conditions (3) and (4) are sufficient to show that for games of the form v_R defined for any $R, S \subseteq N$ by

$$v_R(S) = \begin{cases} 1 & \text{if} \quad R \subseteq S, \\ 0 & \text{if} \quad R \not\subseteq S, \end{cases}$$

and for any nonnegative number c, the value must be

$$\phi_i(cv_R) = \begin{cases} c/r & \text{if} \quad i \in R, \\ 0 & \text{if} \quad i \notin R. \end{cases} \tag{3.1}$$

It can also be shown that any game v can be written as a linear combination of games of the form v_R:

$$v = \textstyle\sum_{R \subseteq N} c_R(v) v_R \tag{3.2}$$

where the coefficients $c_R(v)$ are given by

$$c_R(v) = \sum_{T \subseteq R} (-1)^{r-t} v(T). \tag{3.3}$$

Condition 5 can now be used to demonstrate the remarkable result that the unique function satisfying Conditions 3, 4, and 5 and defined on all games is given by

$$\phi_i(v) = \sum_{S \subseteq N} \left(\frac{(s-1)!(n-s)!}{n!} \right) (v(S) - v(S-i)).$$

4. A UTILITY FUNCTION FOR GAMES

For simplicity of presentation we will, henceforth, confine our attention to the class Z of games which are positive valued, i.e., games for which $v(S) \geq 0$ for all $S \subseteq N$. A position $i \in N$ is called a *dummy*[2] for a game v if it is not contained in every carrier. Denote by $D_i \subseteq Z$ the class of games for which i is a dummy.

It will be convenient to define the games v_0 and v_i given for all $S \subseteq N$ by

$$v_0(S) = 0 \quad \text{and} \quad v_i(S) = \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{if } i \notin S. \end{cases}$$

In the game v_0, all positions are dummies; in v_i all positions but i are dummies.

We will be interested in the mixture set M generated by the set $Z \times N$ of *strategic positions*. Thus, M consists of all lotteries of the form $[p(v, i); (1-p) \times (w, j)]$, where (v, i) and (w, j) are elements of $Z \times N$. We assume that a preference relation $\succcurlyeq$ is defined on M which satisfies Conditions 1 and 2. (Read $(v, i) \succcurlyeq (w, j)$ as "it is preferred to play position i in game v than to play position j in game w.")

The games in the class Z are all defined in terms of some observer's (fixed) utility function u for "simple situations." The preference relation over the set M is assumed to belong to the same observer. We impose the following restrictions on the preferences.

CONDITION 6: For all $i \in N$, $v \in Z$ and for any permutation π, $(v, i) \sim (\pi v, \pi i)$.

CONDITION 7: If $v \in D_i$, then $(v, i) \sim (v_0, i)$, and for every $v \in Z$ and $i \in N$, $(v, i) \succcurlyeq (v_0, i)$ and $(v_i, i) \succ (v_0, i)$.

CONDITION 8: For any number $c > 1$, and for every $v \in Z$, $i \in N$, $(v, i) \sim [(1/c)(cv, i); (1-(1/c))(v_0, i)]$.

Condition 6 merely says that the names of the positions do not determine their desirability in a game. Condition 7 says that playing any position in any game (in the class Z) is at least as desirable as being a dummy in any game, and that there is some strategic position, namely (v_i, i), which is strictly preferable to being a dummy.

[2] Note that this differs slightly from the usual definition of a dummy.

Condition 8 takes note of the fact that games are defined in terms of a utility function. It says that if two strategic positions are identical except for the fact that the utility obtainable in one is a positive multiple of the utility obtainable in the other, then the first is indifferent to the appropriate gamble between the second, and the prospect of receiving zero.

We are now in a position to define a utility function for strategic position, which we shall call *strategic utility* to distinguish it from the utility function u used to define the games. Such a function exists, since the preference $\geqslant$ satisfies Conditions 1 and 2 by assumption.

The *strategic utility* of a game v is the vector $\theta(v) = (\theta_1(v), \theta_2(v), \ldots, \theta_n(v))$, where

$$\theta_i(v) \equiv \theta(v, i) = \frac{p_{ab}((v, i)) - p_{ab}(r_0)}{p_{ab}(r_1) - p_{ab}(r_0)}$$

for probabilities $p_{ab}(\cdot)$ as in (2.3) and for $a, b, r_1, r_0 \in M$ such that $a \geqslant (v, i) \geqslant b$, and $a \geqslant r_1 > r_0 \geqslant b$. Fixing $r_1 = (v_i, i)$ and $r_0 = (v_0, i)$ we get $\theta_i(v_i) = 1$ and $\theta_i(v_0) = 0$. Condition 7 insures that we can always take $b = r_0$, so that $p_{ab}(r_0) = 0$ for all $a \in M$.

We can also prove the following lemmas.

LEMMA 1: *For any permutation π, and for every $i \in N$,*

$$\theta_{\pi i}(\pi v) = \theta_i(v).$$

PROOF: Immediate from the order preserving properties of utility functions, and from Condition 6.

LEMMA 2: *For any number $c \geqslant 0$. and for any $i \in N$,*

$$\theta_i(cv) = c\theta_i(v).$$

PROOF: Without loss of generality, take $c \geqslant 1$.

Case 1: $(cv, i) \geqslant r_1 = (v_i, i)$. Take $a = (cv, i)$ and $b = r_0 = (v_0, i)$. Then

$$\theta_i(cv) = \frac{p_{ab}((cv, i))}{p_{ab}(r_1)} = \frac{1}{p_{ab}(r_1)}.$$

But by Condition 8,

$$(v, i) \sim \left[\frac{1}{c}(cv, i); \left(1 - \frac{1}{c}\right) r_0\right], \quad \text{so} \quad p_{ab}((v, i)) = \frac{1}{c}.$$

Consequently,

$$\theta_i(v) = \frac{p_{ab}((v, i))}{p_{ab}(r_1)} = \frac{1}{c}\theta_i(cv).$$

Case 2: Let $r_1=(v_i, i)\geqslant(cv, i)$. Take $a=r_1$, $b=r_0$. Then $p_{ab}(r_1)=1$, and so $\theta_i(cv)=p_{ab}((cv, i))$. But $(v, i)\sim[(1/c)(cv, i); (1-(1/c))r_0]\sim[(1/c)[p_{ab}((cv, i))a; 1-p_{ab}((cv, i))b]; (1-(1/c)b]$ by definition of $p_{ab}(\cdot)$. But by condition (2.1), this is equal to $[(1/c)p_{ab}((cv, i))a; (1-(1/c)p_{ab}((cv, i)))b]$. So $\theta_i(v)=p_{ab}((v, i))=(1/c)p_{ab}((cv, i))=(1/c)\theta_i(cv)$.

Our procedure so far has been, in effect, to imbed the set of simple prospects in the mixture space of games. Any simple prospect whose utility u is equal to some value $c\geqslant 0$ can be associated with the game cv_i, since

$$\theta_i(cv_i)=c. \tag{4.1}$$

This follows from Lemma 2 and the fact that $\theta_i(v_i)=1$ and $\theta_i(v_0)=0$. By Lemma 1, any position i yields the same result.

Equation (4.1) says that the strategic utility θ is simply an extension of the utility function u which defines the games. The utility function θ is unique (for a fixed u), since (4.1) sets the origin and scale.

In order to evaluate θ for other elements of the set M (i.e., for games not of the form cv_i) we must investigate the risk posture of the preference relation $\geqslant$.

5. RISK POSTURE

We shall distinguish between two kinds of risk. *Ordinary risk* involves the uncertainty which arises from the chance mechanism involved in lotteries, while *strategic risk* involves the uncertainty which arises from the interaction in a game of the strategic players (i.e., those who are not dummies).

The preference $\geqslant$ is *averse to strategic risk* if for every $R\subseteq N$ and all $i\in R$, $(v_i, i)>(rv_R, i)$, where v_R is defined as in Section 3. This means that it is preferable to receive a utility of one for certain (in a game with no other strategic players) than to negotiate how to distribute a utility of r among r players.

If the preference is reversed, we say it is *risk preferring* to strategic risk. The preference relation $\geqslant$ is *neutral* to strategic risk if for all $R\subseteq N$, and every $i\in R$,

$$(v_i, i)\sim(rv_R, i). \tag{5.1}$$

The preference $\geqslant$ is *averse to ordinary risk* if for all $i\in N$, and $v, w\in Z$, $((pw+(1-p)v), i)>[p(w, i); (1-p)(v, i)]$, i.e., if it is preferable to play the game $(pw+(1-p)v)$ than to have a lottery which results in the game w with probability p and the game v with probability $(1-p)$.

Similarly, the preference is *neutral* to ordinary risk if for all games $v, w\in Z$, and for every $i\in N$

$$((pw+(1-p)v), i)\sim[p(w,i); (1-p)(v, i)]. \tag{5.2}$$

We can now prove the following lemmas.

LEMMA 3: *If the preference relation $\succcurlyeq$ is neutral to strategic risk, then*

$$\theta_i(v_R) = \begin{cases} \dfrac{1}{r} & \text{for } i \in R, \\ 0 & \text{for } i \notin R. \end{cases}$$

PROOF: If $i \notin R$, then $\theta_i(v_R) = 0$ by Condition 7 and the fact that $v_R \in D_i$. If $i \in R$, then $\theta_i(v_R) = 1/r$ by (5.1) and Lemma 2.

LEMMA 4: *If the preference relation $\succcurlyeq$ is neutral to ordinary risk, then*

$$\theta(v+w) = \theta(v) + \theta(w),$$

PROOF: For each $i \in N$, $\theta_i(v+w) = \theta_i(2(\frac{1}{2}v + \frac{1}{2}w)) = 2\theta_i(\frac{1}{2}v + \frac{1}{2}w)$ by Lemma 2. But $\theta_i(\frac{1}{2}v + \frac{1}{2}w) = \frac{1}{2}\theta_i(v) + \frac{1}{2}\theta_i(w)$ by (5.2) and (2.2), so $\theta_i(v+w) = \theta_i(v) + \theta_i(w)$.

Thus for a preference relation $\succcurlyeq$ which obeys Conditions 1, 2, 6, 7, and 8, we can prove the following theorem:

THEOREM: *The strategic utility θ is equal to the Shapley value ϕ if and only if the preference relation $\succcurlyeq$ is neutral with respect to both ordinary and strategic risk.*

PROOF: By (3.2) and Lemmas 2 and 4, $\theta(v) = \Sigma_{R \subseteq N} c_R(v)\theta(v_R)$, where the numbers $c_R(v)$ are given by (3.3). But $\theta_i(v_R) = \phi_i(v_R)$ by Lemma 3 and (3.1). So $\theta_i(v) = \Sigma_{R \subseteq N} c_R(v)\phi_i(v_R) = \phi_i(v)$.

6. DISCUSSION

We have shown that the Shapley value is a risk neutral utility function. This fact sheds some light on the axioms which define the value, particularly Conditions 4 and 5. Lemmas 3 and 4 make it clear that these two axioms are intimately related to the two kinds of risk neutrality.

Perhaps it will also serve to illuminate some of the properties of the value. For instance, the relationship between the value and the competitive allocations of a market game might be better understood if conditions could be determined under which preferences tend towards risk neutrality as the competitiveness of the market increases.

It may also prove fruitful to investigate strategic utility functions arising from preferences which are not risk neutral. In particular, the concept of *strategic* risk yields interesting results in this regard. In [**2** and **3**] the author considers some alternate risk postures and characterizes the class of utilities which reflect neutrality to ordinary risk. In this case, the utility of playing a game can be expressed as a function of the attitude towards strategic risk. Preliminary results [**4**] indicate that a similar approach may be fruitful in studying games *without* side payments.

A final comment is in order about the organization of this paper. The normalization of the preferences given by Conditions 6, 7, and 8 seems to be completely natural, regardless of the risk posture which the preference reflects. But, when we are dealing with preferences neutral to ordinary risk, it is not necessary to independently assume Condition 8, because of the following proposition.

PROPOSITION: *Ordinary risk neutrality implies Condition 8.*

PROOF: Let $c>1$, and $p=1/c$. Then by condition (5.2) we have

$$(v_i, i)=\left(\frac{1}{c}cv_i, i\right)=\left(\left(\frac{1}{c}cv_i+\left(1-\frac{1}{c}\right)v_0\right), i\right)$$
$$\sim\left[\frac{1}{c}(cv_i, i);\left(1-\frac{1}{c}\right)(v_0, i)\right].$$

But this is precisely Condition 8, so we are done.

University of Illinois

Manuscript received January, 1976; last revision received September, 1976.

REFERENCES

[1] HERSTEIN, I. N., AND J. MILNOR: "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21 (1953), 291–297.
[2] ROTH, A. E.: "A Note on Values and Multilinear Extensions," *Naval Research Logistics Quarterly*, forthcoming.
[3] ———: "Bargaining Ability, the Utility of Playing a Game, and Models of Coalition Formation," mimeograph, 1976.
[4] ———: "The Nash Solution and the Utility of Bargaining," mimeograph, 1976.
[5] SHAPLEY, L. S.: "A Value for n-Person Games," *Annals of Math Studies*, 28 (1953), 307–317.
[6] VON NEUMANN, J., AND O. MORGENSTERN: *Theory of Games and Economic Behavior.* Princeton: Princeton University Press, 1944, 1947, 1953.