Reprinted from JOURNAL OF MATHEMATICAL PSYCHOLOGY
All Rights Reserved by Academic Press, New York and London

Vol. 16, No. 2, October 1977
Printed in Belgium

Bargaining Ability, the Utility of Playing a Game, and Models of Coalition Formation

ALVIN E. ROTH

Department of Business Administration, University of Illinois, Urbana, Illinois 61801

A utility function for playing a given position in a game is developed as a natural extension of the utility function which defines the rewards available in the game. This function is determined by a player's opinion of his bargaining ability. A characterization of such utilities is obtained which generalizes previous results that the Shapley value and Banzhaf–Coleman index are both cardinal utility functions which reflect different bargaining abilities. The approach taken here is related to models of coalition formation.

1. INTRODUCTION

Considerable literature is devoted to the study of making choices which involve probabilistic uncertainty. In particular, the theory of expected utility (von Neumann and Morgenstern, 1947; Herstein and Milnor, 1953; Fishburn, 1970; Krantz *et al.* 1971) considers the question of how to evaluate, a priori, the relative worth of lotteries, whose outcomes can be described probabilistically. Game theory, on the other hand, is primarily concerned with decisions involving strategic, rather than probabilistic uncertainty. The outcome of a game is uncertain because it depends on the strategic interaction of rational players, which cannot be completely described by probability distributions.

Thus the problem of evaluating, a priori, the relative value of playing in different games appears to be quite different from the problem of establishing a utility for lotteries. The most famous approach to this problem is due to Shapley (1953), and values for restricted classes of games have been proposed by Banzhaf (1968), and Coleman (1971).

However, a recent paper (Roth, 1977a) has shown that the Shapley value is itself a von Neumann–Morgenstern utility function. In fact the Shapley value of a game is an extension of the utility function (for lotteries) used to define the game. In this paper we shall characterize all such extensions. We will show that an individual's utility for playing a particular position in a game is determined by his opinion of his own bargaining ability. We shall also explore the connection which arises between this assessment of bargaining ability and models of coalition formation.

ISSN 0022-2496

2. Definitions

A *game* consists of a finite set of *positions*[1] $N = \{1,..., n\}$ and a superadditive function v from the subsets of N to the real numbers such that $v(\Phi) = 0$. The function v denotes the payoff obtainable by each subset of positions acting together (i.e., by each coalition of players) in units chosen to be linear in the utility function (for lotteries) of a given observer.

Since we shall be interested in comparing different games, we take N to be the common set of positions for all games. Thus n is the maximum number of players who can take part in a game (e.g., n could be the population of the world). In order to distinguish those positions which have an active role in a given game, we define a *carrier* of a game v to be a subset T of N such that for all subsets S of N, $v(S) = v(S \cap T)$. The smallest carrier in a game can be viewed as the set of positions with an active role in the game, since only this set of positions has any effect on the utility obtainable by a coalition. A position which is not in every carrier of a game thus has no effect, and is called a *dummy* position in the game.[2] The smallest carrier will be called the set of *strategic positions* in the game.

If π is a permutation of N (i.e., a one-to-one mapping from N to N) then for all subsets S of N we donate the image of S under π by πS, and define the game πv by $\pi v(\pi S) = v(S)$. We will denote the class of all games by G.

3. Comparing Positions in Games

In order to make comparison between positions in a game and in different games, we shall consider a preference relation defined on the set $N \times G$ of positions in a game. Write $(i, v)\, P(j, w)$ to mean "it is preferable to play position i in game v than to play position j in game w." The letter I will denote indifference, and W will denote weak preference.[3]

We will consider preference relations which are also defined on the mixture set[4] M generated by $N \times G$. That is, preferences are also defined over lotteries whose outcomes are positions in a game. Denote by $[q(i, v);\ (1 - q)\ (j, w)]$ the lottery which with probability q has a player take position i in game v, and with probability $(1 - q)$ take position j in game w. We will henceforth only consider preference relations which have the standard properties of continuity and substitutability[5] on M which insure the existence of an

[1] We speak of "positions" rather than the more customary "players" since we are interested here only in the structural properties of the game. We shall be concerned with the problem of evaluating the different positions from the point of view of a player who must choose among different positions.

[2] This terminology is slightly unconventional; the standard definition of a dummy is a position i such that $v(S \cup i) = v(S) + v(i)$ whenever $i \notin S$.

[3] So aIb means neither aPb nor bPa, and aWb means aPb or aIb.

[4] A mixture set has the properties that for all $a, b \in M [1a; 0b] = a$, $[qa; (1 - q)b] = [(1 - q)b; qa]$, and $[q[pa; (1 - p)b]; (1 - q)b] = [pqa; (1 - pq)b]$ (cf. Fishburn, 1970).

[5] Compare Herstein and Milnor (1953).

expected utility function[6] unique up to an affine transformation. Denote this function by θ, and write $\theta_i(v) = \theta((i, v))$, and $\theta(v) = (\theta_1(v),\ldots, \theta_n(v))$.

The utility for a position in a game is given by[7]

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

where a, b, r_1, and r_0 are elements of M such that $aW(i, v)Wb$, and aWr_1Pr_0Wb, and the numbers $q_{ab}(y)$ are probabilities defined for any y in M such that $aWyWb$ by $yI[q_{ab}(y)a; (1 - q_{ab}(y))b]$. The elements r_1 and r_0 determine the origin and scale, since $\theta(r_1) = 1$, and $\theta(r_0) = 0$.

Since the games themselves are defined in terms of a player's utility for lotteries, we may expect that his preferences for positions in a game will satisfy some additional regularity conditions. That is, his preferences for positions in games, and the resulting utility function should be consistent with his utility function for lotteries.

It will be convenient to define, for each position i, the game v_i given by

$$\begin{aligned} v_i(S) &= 1 \qquad \text{if } i \in S, \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

All positions other than i are dummies in games of the form cv_i, so that the player in position i may be sure of getting a utility of c. (This observation will provide the appropriate normalization for the utility θ.) Denote by v_0 the game in which all players are dummies, (i.e., the game $v_0(S) = 0$), and let D_i be the class of games in which position i is a dummy.

The regularity conditions which we impose on the preferences are as follows.

R1. For all $i \in N$, $v \in G$ and for any permutation π, $(i, v)\, I(\pi i, \pi v)$. This condition says simply that the *names* of the positions do not affect their desirability. An immediate consequence is the following.

LEMMA 1. $\theta_{\pi i}(\pi v) = \theta_i(v)$.

The second regularity condition is

R2. If $v \in D_i$, then $(i, v)\, I(i, v_0)$. Also, $(i, v_i)\, P(i, v_0)$.

This condition says that being a dummy in a game is not preferable to being a dummy in the game v_0, and that the position (i, v_i) is preferable to playing a dummy position.

The last regularity condition is considerably stronger, but follows naturally from the fact that the games are defined in terms of an expected utility function.

R3. $(i, (qw + (1 - q)v))\, I[q(i, w); (1 - q)(i, v)]$.

[6] An expected utility function on a mixture space has the property that $u([qa; (1 - q)b]) = qu(a) + (1 - q)\, u(b)$. That is, the utility of a lottery is its expected utility.

[7] Compare Herstein and Milnor.

This says that a player is indifferent between playing position i in the game $(qw + (1 - q)v)$ or having the appropriate lottery between the games w and v. In Roth (1977a) this condition is called *ordinary risk neutrality*, since it specifies the preferences over lotteries.

The next sections will study the class of utility functions for positions in games which arise from preferences obeying the above conditions. Section 4 will study properties shared by the entire class of utilities, subject to the natural normalization that $\theta_i(v_i) = 1$, and $\theta_i(v_0) = 0$. Section 5 will show that a specification of an individual's beliefs about his bargaining ability uniquely determines his utility for a position in a game.

4. Extended Utility Functions

A utility function on positions in games which reflects the conditions of the previous section will be called an *extended utility function*. This terminology is justified by the following theorem.

THEOREM 1. *For any* $c \geqslant 0$, *and any* $(i, v) \in N \times G$, $\theta_i(cv) = c\theta_i(v)$.

Proof. Without loss of generality, take $c \geqslant 1$. Note that $(i, v) = (i, ((1/c)cv + (1 - (1/c))v_0))\ I[(1/c)(i, cv); (1 - (1/c))\ (i, v_0)]$ by R3. So $\theta_i(v) = (1/c)\theta_i(cv) + (1 - (1/c))\ \theta_i(v_0) = (1/c)\theta_i(cv)$. Q.E.D.

In particular, $\theta_i(cv_i) = c$, so the utility function θ can be regarded as an extension of the utility function in terms of which the games are defined. The following theorem also holds for extended utility functions.

THEOREM 2. *For any* $v, w \in G$, $\theta(v + w) = \theta(v) + \theta(w)$.

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 Theorem 1. But by R3, $\theta_i(\frac{1}{2}v + \frac{1}{2}w) = \theta_i([\frac{1}{2}v; \frac{1}{2}w1) = \frac{1}{2}\theta_i(v) + \frac{1}{2}\theta_i(w)$, since θ is an expected utility function. So $\theta_i(v + w) = \theta_i(v) + \theta_i(w)$.

5. Strategic Risk Posture

Any game with more than one strategic position involves some potential uncertainty as to the outcome, arising from the interaction of the strategic players. This kind of uncertainty we call strategic risk. In order to describe a given player's preferences for situations involving strategic risk, it will be convenient to consider for each subset R of N, the game v_R defined by

$$v_R(S) = 1 \quad \text{if } R \subset S,$$
$$= 0 \quad \text{otherwise.}$$

A game of the form v_R is essentially the simplest game that can be played among r strategic players.[8]

To measure a player's opinion of his own bargaining ability, consider how he evaluates the prospect of being a strategic player in a game of the form v_R. The prospect will be more desirable if he thinks he is a strong bargainer than if he thinks he is a weak bargainer.

Define the *certain equivalent* of a strategic position in a game v_R to be the number $f(r)$ such that the prospect of receiving $f(r)$ for certain is exactly as desirable as the prospect of playing the strategic position.[9] That is, $f(r)$ is the number such that (for $i \in R$)

$$(i, v_R)\, I(i, f(r)v_i).$$

Note that $f(1) = 1$.

Using the terminology of Roth (1977a), we say that the preference is *neutral* to strategic risk if $f(r) = 1/r$ for $r = 1,..., n$. The preference is strategic risk *averse* if $f(r) < 1/r$, and strategic risk *preferring* if $f(r) > 1/r$. The utility of playing a position in a game v_R is given by the following lemma.

LEMMA 2.

$$\begin{aligned} \theta_i(v_R) &= f(r) && \textit{if } i \in R, \\ &= 0 && \textit{otherwise.} \end{aligned}$$

Proof. If $i \notin R$, then $v_R \in D_i$ and $\theta_i(v_R) = \theta_i(v_0) = 0$, by R2. If $i \in R$, then $\theta_i(v_R) = \theta_i(f(r)v_i) = f(r)\,\theta_i(v_i) = f(r)$, by Theorem 1.

It turns out that an extended utility function is completely determined by the numbers $f(r)$. In fact, we have the following

REPRESENTATION THEOREM. *An extended utility function has the form*

$$\theta_i(v) = \sum_{T \subset N} k(t)[v(T) - v(T - i)] \tag{1}$$

where

$$k(t) = \sum_{r=t}^{n} (-1)^{r-t} \binom{n-t}{r-t} f(r).$$

Proof. Every game v is a sum of games of the form v_R. In fact, (cf. Shapley, 1953), $v = \sum_{R \subset N} c_R v_R$ where $c_R = \sum_{T \subset R} (-1)^{r-t}\, v(T)$. By Theorems 1 and 2,

$$\theta_i(v) = \sum_{R \subset N} c_R \theta_i(v_R) = \sum_{\substack{R \subset N \\ i \in R}} c_R f(r) = \sum_{\substack{R \subset N \\ i \in R}} \sum_{T \subset R} (-1)^{r-t}\, v(T) f(r).$$

[8] The cardinality of sets R, S, T is denote r, s, t.

[9] We take the point of view that a player does not know *who* will occupy the other positions in a game. Consequently, his certain equivalent for a game v_R depends only on r.

Reversing the order of summation, $\theta_i(v) = \sum_{T\subset N} \{\sum_{R\subset N, R\supset \{T\cup i\}} (-1)^{r-t} f(r)\} v(T)$. If we denote the term in brackets by $g_i(T)$, then we note that $g_i(T) = -g_i(T-i)$ when $i \in T$. So $\theta_i(v) = \sum_{T\subset N, i\in T} g_i(T)[v(T) - v(T-i)]$. But there are $\binom{n-t}{r-t}$ coalitions of size r which contain T, so $g_i(T) = \sum_{r=t}^{n} (-1)^{r-t}\binom{n-t}{r-t} f(r) = k(t)$. Since $[v(T) - v(T-i)] = 0$ unless $i \in T$, we are done.

As mentioned earlier, one function which has been extensively studied is the Shapley value $\phi = (\phi_1, ..., \phi_n)$, given by

$$\phi_i(v) = \sum_{S\subset N} \frac{(s-1)!\,(n-s)!}{n!} [v(S) - v(S-i)].$$

This is the extended utility function corresponding to a posture of strategic risk neutrality. More formally, we have the following corollary.

COROLLARY 1 (Roth, 1977a). *If $f(r) = 1/r$ then the extended utility function equals the Shapley value.*

Another function which has received some attention in the literature (cf. Owen, 1975; Blair, 1975; Dubey, 1975; Roth, 1977b, c)[10] is the Banzhaf–Coleman index $b = (b_1, ..., b_n)$ given by

$$b_i(v) = \sum_{S\subset N} 1/(2^{n-1})[v(S) - v(S-i)].$$

The Banzhaf–Coleman index is an extended utility function reflecting preferences averse to strategic risk.

COROLLARY 2 (Roth, 1977b). *If $f(r) = 1/(2^{r-1})$ then the extended utility function equals the Banzhaf–Coleman index.*

Thus the Banzhaf–Coleman index is a utility function in which a player's utility for a strategic position in a game v_R is inversely proportional to the number of ways the $r-1$ other strategic players can form coaliations.

6. Probabilistic Models of Coalition Formation

Many of the experimental studies of coalition formation have concentrated on the class of *simple* games. A game v is *simple* if for all subsets S of N, $v(S) = 0$ or 1, and if S is contained in T, then $v(S) = 1$ implies $v(T) = 1$. A simple game can be characterized by the class W of *winning coalitions*, defined by $W = \{S \subset N \mid v(S) = 1\}$. Position i is *critical* in a coalition S if S is winning, but $S - i$ is not.

Under some circumstances it may be possible to assess the probability that a given winning coalition will form. Komorita (1974) argues that this probability is a function of the coalition size. In this case, the probability that a player in position i will be a critical

[10] Some of these papers are concerned primarily with the class of *simple* games (see Section 6).

member of a winning coalition is given by $P_i = \sum_{T \in W, T-i \notin W} P(t)$, where $P(t)$ is the probability that a coalition of size t will form. But $\sum_{T \in W, T-i \notin W} P(t) = \sum_{T \subset N} P(t)[v(T) - v(T-i)]$, since $[v(T) - v(T-i)]$ equals one if position i is critical in T, and equals zero otherwise.

If player i expect to receive a reward of $r(t)$ when he is a critical member of a winning coalition of size t (and a reward of zero otherwise), then his expected reward is equal to $\sum_{T \subset N} r(t)\, P(t)[v(T) - v(T-i)]$. Thus the expected reward has the form of an extended utility function. Conversely, the Shapley value and Banzhaf–Coleman index are often interpreted as expected reward functions in which $r(t) = 1$.

7. Conclusion

In Sections 1–5 we consider how to determine the utility of playing a position in a game. The derivation of this utility function depends on comparisons between different games. In particular, a player's extended utility function is uniquely determined by his opinion of his bargaining ability, as expressed in his evaluation of games of the form v_R.

In Section 6, motivated by the experimental work of Komorita and others,[11] we considered a probabilistic model of coalition formation. An expression for the expected reward of a player in a given position of a particular game was given. Rather than depending on a comparison between different games, this expression depends on detailed consideration of the game in question. In particular, the expected reward depends on an assessment both of how well the player will do if a given coalition forms, as well as the probability that the coalition will form.

As noted, the expressions for the expected reward and for the extended utility have the same functional form. The connection is more than a purely formal one, since a player's experiences in a game (which influence his assessment of the expected reward) should influence his opinion of his bargaining ability, while his opinion of his bargaining ability should influence his expectations of reward in a given game.

References

Banzhaf, J. F., III. One man 3.312 votes: A mathematical analysis of the electoral college. *Villanova Law Review*, 1968, 13.

Blair, D. H. *Power valuations of finite voting games*. Presented at Econometric Society Winter Meeting, 1975.

Coleman, J. S. Control of collectivities and the power of a collectivity to act. In B. Lieberman (Ed.), *Social Choice*. New York: Gordon & Breach, 1971.

Dubey, P. *Some results on values of finite and infinite games*. Ph. D. thesis, Cornell University, 1975.

Fishburn, P. C. *Utility theory for decision making*. New York: Wiley, 1970.

Herstein, I. N., & Milnor, J. An axiomatic approach to measurable utility. *Econometrica*, 1953, 21.

Komorita, S. S. A weighted probability model of coalition formation. *Psychological Review*, 1974, 81.

[11] Compare Murnighan *et al.*

KRANTZ, D. H., LUCE, R. D., SUPPES, P., & TVERSKY, A. *Foundations of measurement volume I.* New York: Academic Press, 1971.

MURNIGHAN, J. K., KOMORITA, S. S., & SWAJKOWSKI, E. Theories of coalition formation and the effects of reference groups. *Journal of Experimental Social Psychology*, 1977, 13.

OWEN, G. Multilinear extensions and the Banzhaf value. *Naval Research Logistics Quarterly*, 1975, **22.**

ROTH, A. E. The Shapley value as a von Neumann–Morgenstern utility. *Econometrica*, 1977a, **45.**

ROTH, A. E. A note on values and multilinear extensions. *Naval Research Logistics Quarterly*, 1977b, forthcoming.

ROTH, A. E. Utility functions for simple games. *Journal of Economic Theory*, 1977c, forthcoming.

SHAPLEY, L. S. A value for *n*-person games. *Annals of Mathematics Study*, 1953, 28.

VON NEUMANN, J., & MORGENSTERN, O. *Theory of games and economic behavior*. Princeton: Princeton Univ. Press, 1944, 1947, 1953.

RECEIVED: July 12, 1976

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.