RESEARCH ANNOUNCEMENTS

The purpose of this department is to provide early announcement of significant new results, with some indications of proof. Although ordinarily a research announcement should be a brief summary of a paper to be published in full elsewhere, papers giving complete proofs of results of exceptional interest are also solicited.

VON NEUMANN-MORGENSTERN SOLUTIONS TO CO-OPERATIVE GAMES WITHOUT SIDE PAYMENTS

BY R. J. AUMANN AND B. PELEG

Communicated by A. W. Tucker, January 7, 1960

The use of side payments in the classical[1] theory of n-person games involves three restrictive assumptions. First, there must be a common medium of exchange (such as money) in which the side payments may be effected; next, the side payments must be physically and legally feasible; and finally, it is assumed that utility is "unrestrictedly transferable," i.e. that each player's utility for money[2] is a linear function of the amount of money.[3] These assumptions severely limit the applicability of the classical theory; in particular, the last assumption has been characterized by Luce and Raiffa [2, p. 233] as being "exceedingly restrictive—for many purposes it renders n-person theory next to useless." It is the purpose of this paper to present the outline of a theory that parallels the classical theory, but *makes no use of side payments*.[4] Our definitions are related to those given in [2, p. 234] and in [3], but whereas the previous work went no further than proposing definitions, the theory outlined here contains results which generalize a considerable portion of the classical theory. It thus demonstrates that the restrictive side payment assumption is not necessary for the development of a theory based on the ideas of von Neumann and Morgenstern. Only a general description of the theory and statements of the more important theorems will be included here; details and proofs will be published elsewhere.

[1] We will use the word "classical" to denote the von Neumann-Morgenstern theory as described in [1] and in [4].

[2] Or any other medium of exchange.

[3] See [2, p. 168]. It can be proved that when $n \geqq 3$, linearity of the utilities in money is necessary and sufficient for the existence of an unrestrictedly transferable utility.

[4] In particular, our theory is of course also applicable to the case in which side payments are permitted. When, in addition, utility is unrestrictedly transferable, then our theory reduces to the classical theory.

1. **Effectiveness.** Let us fix attention on a given finite n-person game, and let N denote the set of players. Let E^N denote an n-dimensional euclidean space, and let us index the coordinates of points in E^N by the members of N. The points of E^N will be called *payoff vectors*; if $x \in E^N$ and $i \in N$, then x_i will denote the coordinate of x corresponding to player i, and will be called the *payoff* to i.

Intuitively, a coalition B is *effective* for a payoff vector x if the members of B, by joining forces, can play so that each player i in B receives at least x_i. This intuitive definition is open to a number of interpretations. The rather conservative one adopted by von Neumann and Morgenstern assumes that the most the members of B can count on is what they can get if the players of $N-B$ form a coalition whose purpose it is to minimize the payoff to B. There are at least two generalizations of this notion of effectiveness to the case in which there are no side payments:

(i) *A coalition B is said to be α-effective for the payoff vector x if there is a strategy*[5] *for B, such that for each strategy used by $N-B$, each member i of B receives at least x_i.*

(ii) *A coalition B is said to be β-effective for the payoff vector x, if for each strategy used by $N-B$, there is a strategy for B such that each member i of B receives at least x_i.*

Roughly, α-effectiveness means that B can assure itself of its portion of x independently of the actions of $N-B$, whereas β-effectiveness means that $N-B$ cannot prevent B from obtaining its (B's) portion of x. In the classical theory the two notions are equivalent, but this is not the case when side payments are forbidden. Which of the two definitions is preferable is a matter of taste; both have appeared, in more or less disguised form, in the previous literature [**2**, p. 175; **3**; **5**]. There seems to be a tendency to consider α-effectiveness as intuitively more appealing; on the other hand, there is evidence that β-effectiveness may eventually turn out to be the more significant concept.[6] The present theory applies equally well to both notions.

2. **Axiomatic treatment.** It is possible to define many of the basic notions of n-person theory—domination, solution, core, etc.—in terms of effectiveness. Of course the objects we will get will usually depend on what kind of effectiveness we started with. Thus we will define the α-core and the β-core, but for a given game they usually

[5] The word "strategy" as used in this paper means what has been variously called "correlated mixed strategy" [**2**, p. 116], "joint randomized strategy" [**2**, p. 116], "cooperative strategy" [**2**, p. 175], and "correlated strategy B-vector" [**5**].

[6] See §6.

differ; similarly with α-solutions and β-solutions, etc. Nevertheless, it is possible to prove a considerable number of general theorems which hold for either kind of effectiveness; the proofs of these theorems make use only of certain basic properties common to both kinds. The situation invites axiomatic treatment.

An n-person "characteristic function" is a set N with n members, together with a function v that carries each subset B of N into a subset $v(B)$ of E^N so that

(1) *$v(B)$ is convex;*

(2) *$v(B)$ is closed;*

(3) $v(\emptyset)=E^N$;[7]

(4) *if $x\in v(B)$, $y\in E^N$, and for all $i\in B$, $y_i\leqq x_i$, then $y\in v(B)$; and*

(5) *if B_1 and B_2 are disjoint, then $v(B_1\cup B_2)\supset v(B_1)\cap v(B_2)$.*

An n-person "game" is an n-person characteristic function (N, v) together with a convex compact polyhedral subset H of $v(N)$.

An n-person game as just defined actually represents more than a game in the usual sense; it is a game together with a concept of effectiveness. The set H is the set of all "feasible" or "attainable" payoff vectors, i.e. the set of all payoff vectors which can be attained by a joint strategy of N. $v(B)$ represents the set of all payoff vectors for which B is effective. Conditions (1), (2), and (3) are self-explanatory. Condition (4) says that if a coalition B is effective for a payoff vector x, then it is also effective for any payoff vector with smaller (or equal) payoffs to its members. Condition (5) is the natural generalization of super-additivity of the characteristic function in the classical theory.[8]

In order to justify these definitions, it is necessary to show that an arbitrary finite game, when combined with the concept either of α-effectiveness or of β-effectiveness, satisfies our definition of a game.[9] For the most part this is straightforward; the only deep part occurs in verifying condition (5) in the case of β-effectiveness, where use is made of Kakutani's fixed point theorem.

One of the chief advantages of the axiomatic approach is its flexibility: it can be used not only with the notions of effectiveness described in §1, which are based on the conservative approach that characterizes the classical theory, but also with many other notions of effectiveness. For example, we may prefer an effectiveness notion based on the ideas of ψ-stability [2, pp. 163–168, 174–176, 220–236],

---

[7] $\emptyset$ denotes the empty set.

[8] Condition (5) is not needed for any of the results stated in this paper. It was included in order to underscore the parallelism with the classical theory, and with the hope that stronger axioms will eventually yield a richer theory.

[9] Where $v(B)$ and H have the meanings described in the previous paragraph.

Such notions can be constructed in a number of ways; the general idea would be that in order for a coalition B to be effective for a pair consisting of a payoff vector x and a coalition structure τ, the coalition B must be attainable from the given coalition structure τ, and no attainable combination of coalitions in $N-B$ should be able to prevent B from obtaining its portion of the payoff vector x. Specifically, we would say that B is effective for (x, τ) if $B\in\psi(\tau)$ and there is a strategy for B such that for each partition $(B_1, \cdots, B_k)$ of $N-B$ into members of $\psi(\tau)$, and each k-tuple of strategies used by the B_j, each member i of B receives at least x_i. A related but different notion can be obtained if we reverse the quantifiers. For fixed τ these two notions of effectiveness satisfy all the axioms except (5), and therefore the results stated in this paper hold for them as well (cf. footnote 8).

3. **Domination and solution.** Fix an n-person game $G=(N, v, H)$. A payoff vector x is said to *dominate* a payoff vector y *via* B if $x\in v(B)$ and $x_i>y_i$ for all $i\in B$; x is said to *dominate* y if there is a B such that x dominates y via B. If K is an arbitrary set of payoff vectors, we define dom K to be the set of all payoff vectors dominated by at least one member of K. If P is an arbitrary set of payoff vectors, then a subset K of P is said to be *P-stable* if $K\cap\text{dom } K$ is empty and $K\cup\text{dom } K\supset P$. The set $P-\text{dom } P$ is called the *P-core*. A payoff vector x is said to *majorize* a payoff vector y if x dominates y and all z that dominate x also dominate y. All the lemmas and theorems of §1 of [**4**] concerning domination, P-stability, the P-core and majorization remain true in this context; the proofs go through essentially unchanged.

It is easy to show that for each $i\in N$, there is an extended real number[10] v_i such that $v(\{i\})=\{x\in E^N: x_i\leqq v_i\}$. A payoff vector x is called *individually rational* if $x_i\geqq v_i$ for each $i\in N$. x is called *group rational* if there is no $y\in H$ such that $y_i>x_i$ for each $i\in N$. Let us denote by $\overline{A}$ the set of individually rational members of H, and by A the set of members of $\overline{A}$ that are also group rational. Then it can be proved that a subset of H is A-stable if and only if it is $\overline{A}$-stable. This justifies us in defining a *solution* of G to be an A-stable set.[11]

The next step is to investigate which games are solvable and what their solutions are. First of all, it is easy to show that all 2-person games have a unique solution,[12] namely all of A. We next investigate

[10] A real number or $+\infty$ or $-\infty$.

[11] If K is a solution of G, we will also say that K *solves* G, and that G is *solvable*.

[12] This solution is closely related to the *negotiation set* of a 2-person cooperative game [**2**, p. 118], but is *not* the same thing.

3-person zero-sum[13] games. Unlike the situation in the classical theory, we now find a large number of essentially different games, whose solutions exhibit the greatest variety and complexity. The basic theorem is

THEOREM 1. *Every 3-person zero-sum game is solvable.*

The proof, which is rather involved, proceeds by dividing A into regions, solving each region separately, and then combining the regional solutions into a solution for all of A. The shapes of the regions depend on the $v(B)$ and on their interrelationships. The question of the solvability of 3-person general-sum or 4-person zero-sum games remains open.

Incidentally, Theorem 1 is the only one of our results for which the assumption that the $v(B)$ be convex (condition (1)) is required.[14]

4. Composition. Let $G_1=(N_1, v_1, H_1)$ and $G_2=(N_2, v_2, H_2)$ be games whose player sets N_1 and N_2 are disjoint. Intuitively, the *composition* G of G_1 and G_2 is the game each play of which consists of a play of G_1 and a play of G_2, played without any interconnection. Formally, we define[15] $G=(N, v, H)$, where $N=N_1\cup N_2$, $H=H_1\times H_2$, and for each $B\subset N$, $v(B)=v_1(B\cap N_1)\times v_2(B\cap N_2)$.

THEOREM 2. *A necessary and sufficient condition that a subset K of H solve G is that it be of the form $K_1\times K_2$, where K_1 solves G_1 and K_2 solves G_2.*

The simplicity of this result is somewhat surprising, in view of the fact that the corresponding result in the classical theory is much more complicated. The complexity of the classical result is explained by the fact that it permits side payments between members of N_1 and members of N_2, whereas no such intercourse can be possible in our framework; thus although the classical theory is a special case of our theory, the composition of two games in the classical sense yields a game which is in general not the composition of the games in our sense. All of our solutions appear in the classical theory, but the converse is not true. Our solutions are precisely those in which no "tribute" is paid by either group of players (cf. [1, §46.11.2, p. 401]).

5. The core.

THEOREM 3. *The A-core and the $\bar{A}$-core coincide.*

[13] A game is said to be *zero-sum* if H is contained in the plane $\sum_{i\in N} x_i=0$.

[14] The question as to whether the theorem holds without this assumption remains open; certainly the proof does not go through.

[15] $\times$ denotes cartesian product.

This theorem justifies us in defining the *core* of a game to be its A-core. The proof makes essential use of the fact that H is polyhedral (this is the only theorem for which this assumption is needed); indeed, if this assumption is dropped, the theorem becomes false.

Shapley [6] has conjectured that in the classical theory, the intersection of all solutions is the core. This is not true in our theory; indeed, there is a 3-person zero-sum game with a unique solution, which strictly includes the core.[16]

6. The β-core and the supergame. The *supergame* of a game[17] Γ is the game each play of which consists of an infinite sequence of plays of Γ. A *strong equilibrium point* in an n-person game [5] is, roughly speaking, an n-tuple $\{\xi_i\}_{i\in N}$ of strategies with the property that if the members j of any coalition B use strategies different from the ξ_j, while the players not in B keep using the ξ_i, then at least one player in B will not profit from the change, i.e. will get no more than he would have gotten had all the players used the ξ_i. Strong equilibrium points are strengthened forms of the Nash equilibrium points [7]; at a Nash equilibrium point there is no direct incentive for any *individual* to change his strategy, whereas at a strong equilibrium point there is no direct incentive for any *coalition* to change its strategy.

In [5] the concept of a c-acceptable payoff vector is defined, and it is shown that a payoff vector is c-acceptable in a given finite game if and only if it is the vector of payoffs to a strong equilibrium point in the corresponding supergame. It can be shown that the set of c-acceptable payoff vectors coincides with the β-core.[18] Hence *the β-core of a finite game is precisely the set of payoff vectors to strong equilibrium points in the corresponding supergame.*

7. The "extended" theory. The definition of an *extended game* is similar to that of a game, with the single exception that H is not assumed to be a subset of $v(N)$ but merely of E^N. In the classical theory extended games are important as a theoretical tool in composition theory; they were first considered by von Neumann and Morgenstern [1, Chapter X].

Theorems 1 and 2 remain true as they stand for extended games. Theorem 3 must be adjusted to read "The A-core is the intersection

[16] This solution is disconnected, so that it also provides a counter-example to another conjecture of Shapley, namely that the union of all solutions is connected. Of course this example does not yet settle these questions for the classical theory.

[17] We are now referring to "game" in the ordinary sense of the word, not that defined in §2.

[18] I.e. the core if we use β-effectiveness as our definition of effectiveness.

of the $\bar{A}$-core with A." §1 of [**4**] generalizes as before; however, we have not succeeded in obtaining a relation between the A-stable sets and the $\bar{A}$-stable sets in extended games.

REFERENCES

1. J. von Neumann and O. Morgenstern, *Theory of games and economic behavior*, Princeton University Press, 1943, 2nd ed., 1947.

2. R. D. Luce and H. Raiffa, *Games and decisions*, John Wiley, 1957.

3. L. S. Shapley and M. Shubik, *Solutions of n-person games with ordinal utilities* (abstract), Econometrica vol. 21 (1953) p. 348.

4. D. B. Gillies, *Solutions to general non-zero-sum games*, Contributions to the Theory of Games IV, Princeton University Press, 1959, pp. 47–85.

5. R. J. Aumann, *Acceptable points in general cooperative n-person games*, ibid., pp. 287–324.

6. L. S. Shapley, *Open questions* (dittoed), Report of an Informal Conference on the Theory of n-Person Games held at Princeton University, March 20–21, 1953, p. 15.

7. J. F. Nash, *Non-cooperative games*, Ann. of Math. vol. 54 (1951) pp. 286–295.

THE HEBREW UNIVERSITY,
JERUSALEM, ISRAEL