DISCUSSION PAPER NO. 203

STRAIGHTFORWARDNESS OF GAME FORMS WITH LOTTERIES AS OUTCOMES

by

ALLAN GIBBARD

Preliminary version, February, 1976

Department of Philosophy

University of Pittsburgh

ABSTRACT: Where _alternatives_, _players_, and _strategies_ for each player are finitely many, a _game form_ assigns a lottery over alternatives to each configuration of players' strategies. It is _straightforward_ iff it guarantees that each player, whatever his utilities, will have a dominant strategy. It is _unilateral_ iff only one player can influence the outcome, and _duple_ iff it restricts the final outcome to a fixed pair of alternatives. Any straightforward game form, it is shown, is, on a domain which gives each player a dominant strategy for each utility scale, a probability mixture of game forms, each unilateral or duple.

1. INTRODUCTION

A *game form* is a system which makes an outcome depend on individual actions of some kind, called *strategies*. It is thus a "game" in the sense of von Neumann and Morgenstern (1947).

A game form is *determinate* if the dependence of the outcome on individual strategies involves no element of chance. A player's strategy is *dominant* for him with respect to a weak ordering of the alternatives iff no matter what anyone else does, the strategy secures an outcome at least as high in that weak ordering as is any other lottery he can secure given the strategies of others. A determinate game form is *straightforward* iff each player, for each weak ordering of the alternatives, has a strategy which is dominant with respect to that weak ordering. The only straightforward determinate game forms are trivial: a straightforward determinate game form either restricts the attainable outcomes in advance to no more than two, or makes one player a dictator among attainable outcomes. (This is shown in Gibbard, 1973).

This paper deals with game forms of a more general kind: systems which make an outcome depend on individual strategies in a way that may involve chance. The systems to be considered have finitely many alternatives, finitely many players, and finitely many strategies for each player. A *game form* in this expanded sense assigns to each configuration of individual strategies, or *strategy profiles*, a lottery among alternatives.

Whether a game form makes a given strategy dominant now depends on a player's preferences among lotteries: a strategy is <u>dominant</u> for a player <u>with respect to</u> a weak ordering of all lotteries over alternatives iff no matter what anyone else does, the strategy secures a lottery at least as high in that weak ordering as is any other lottery which he could secure given the strategies of others. A weak ordering of lotteries is <u>coherent</u> iff it satisfies standard conditions, such as those of von Neumann and Morgenstern (1947, p. 26), which entail that it is an ordering by expected utility on some cardinal scale; only those weak orderings of lotteries which are coherent will concern us here. A game form is <u>straightforward</u> iff each player, for each coherent weak ordering of all lotteries over the alternatives, has a strategy which is dominant with respect to that weak ordering.

Any straightforward game form, it will be shown, is, on a slightly restricted domain of individual strategies, a probability mixture of game forms each of which either accords a single player a monopoly of influence, or restricts the final outcome to a fixed pair of alternatives. Game forms of the first kind will be called <u>unilateral</u>, and of the second kind, <u>duple</u>.

The contents of this theorem are perhaps best elucidated through its application to game forms of a special kind, called "decision schemes". A <u>decision scheme</u> is a game form in which players vote weak orderings of the alternatives.[1] The strategy set of each player, then, consists of all weak orderings of the alternatives; as a game form, a decision scheme assigns probabilities to the alternatives on the basis of the way people vote. The theorem in this paper is a generalization of a theorem about decision schemes (in Gibbard, 1976).[1] We can think of the weak ordering a player votes as truly or falsely representing his preferences among the alternatives, and ask what kinds of decision schemes, if any, logically guarantee that no player will ever benefit from misrepresenting his preferences. A decision scheme will be called _strategy-proof_ iff it logically guarantees that honest voting is always a dominant strategy. Now whether a strategy is dominant depends on a voter's preferences among various lotteries the scheme may yield. Honest voting consists in writing down the weak ordering of alternatives as sure things which, in the obvious sense, _fits_ one's weak ordering of all lotteries over alternatives. A decision scheme is _strategy-proof_, then, iff for any voter and for any coherent ordering P^* of all lotteries over alternatives, voting the weak ordering of the alternatives which fits P^* is a dominant strategy with respect to P^*. (A strategy-proof decision scheme is thus straightforward).

There indeed are strategy-proof decision schemes: one is a random dictatorship,[2] a second is pairwise majority rule over a random pair, and a third is a system which chooses randomly between the first two. (These are discussed in Gibbard, 1976). The theorem proved in that paper is that all strategy-proof decision schemes resemble the ones that were just described. In the senses given earlier to the terms 'unilateral' and 'duple', a random dictatorship is a probability mixture of decision schemes which are unilateral; pairwise majority rule over a random pair is a probability mixture of decision schemes which are duple, and a random choice between these two systems yields a probability mixture of decision schemes of the two kinds. The earlier theorem, roughly put, is that any strategy-proof decision scheme is a probability mixture of decision schemes, each of which is either unilateral or duple. This rough statement, though, needs an important qualification: the statement holds in general only for a restricted domain, on which voters do not express indifference between alternatives. A correct statement of the theorem is this: If a decision scheme is strategy-proof, then on a domain obtained by restricting voters' ballots to strong orderings, the scheme is a probability mixture of decision schemes, each of which is either unilateral or duple.[3]

This result is generalized by the theorem to be proved in this paper. A set of strategies is <u>adequate</u> for a player iff for any coherent weak ordering of all lotteries over the alternatives, the set includes a strategy which is dominant for the player with respect to that weak ordering. In a weakened form, the theorem in this paper is as follows: If a game form g is straightforward, then each player i can be assigned an adequate subset S_i^* of his strategy set in such a way that when the strategies of each player i are restricted to the set S^*_i, g is a probability mixture of game forms each of which is unilateral or duple. Now take a strategy-proof decision scheme d. Since d is strategy-proof, it is straightforward. It can be shown, moreover, that the set of all strong orderings of the alternatives constitutes a set of strategies which is adequate. The theorem on decision schemes says that when the strategy set for each player is restricted to strong orderings of the alternatives, d is a probability mixture of unilateral and duple decision schemes.

The theorem in this paper also specifies an assignment of a strategy subset S^*_i to each player i that will do the required job. A coherent weak ordering of lotteries can be represented by a cardinal utility scale, which is an assignment of a real number to each alternative. Such an assignment can be regarded as a vector in a space with a dimension for each alternative. For each player and strategy, let its <u>domain of dominance</u> be the set of all utility vectors with respect to which it is dominant. We can ask whether the domain of dominance of a strategy has interior points; if it does, call the strategy <u>versatile</u>. The theorem in this paper in its full form is that where g is a straightforward game form, (i) for each player, the set of versatile strategies is adequate, and (ii) with each player restricted to versatile strategies, g is a probability mixture of game forms, each of which is either unilateral or duple.

Now if a decision scheme is strategy-proof, then every strong ordering is versatile. For its domain of dominance, by the definition of strategy-proofness, includes every utility scale which it fits: where P orders the alternatives $x_m x_{m-1} \ldots x_2 x_1$, its domain of dominance includes every utility scale U such that

$$U(x_m) > U(x_{m-1}) > \ldots > U(x_2) > U(x_1).$$

This set of utility scales has interior points; an example is the scale U^* such that $U^*(x_\iota) = \iota$ for each $\iota (1 \leq \iota \leq m)$. From this fact and the theorem in this paper, then, the theorem on strategy-proof decision schemes follows: A strategy-proof decision scheme is, when the strategy set of each player is restricted to strong orderings, a probability mixture of unilateral and duple decision schemes.

The theorem on straightforward game forms has two corollaries which may help to clarify its import; the corollaries tell how requirements beyond straightforwardness yield an even more restricted class of game forms. One property we might desire in a game form is that if everyone likes the same alternative best, they shoul be able jointly to bring about its selection. For each alternative, then, there should be a strategy profile which results in that alternative as a sure thing. A straightforward game form which satisfies this requirement, the first corollary says, is, with each player restricted to his versatile strategies, a probability mixture of game forms which are dictatorial, where a game form is dictatorial iff there is a player who, for each alternative, has a strategy which independently of what anyone else does produces that alternative as a sure thing.[4]

A further property that might be demanded of a game form is this. Assign each player i a weak ordering P_i of all lotteries over the alternatives. A lottery is Pareto optimal ex ante with respect to that assignment iff there is no other lottery which ranks higher in everyone's ordering.[5] We might require of a game form that for each assignment of a coherent weak ordering of all lotteries to each player, the game form yield, for some strategy profile, a lottery which is Pareto optimal ex ante with respect to that assignment. A straightforward game form that satisfies this requirement is itself dictatorial; that is the second corollary. These corollaries are formally stated in Section 3.

The nonstraightforwardness of certain voting procedures was studied by Farquharson (1969), who introduced the term 'straightforward' in a somewhat different sense from that used here. Zeckhauser (1969, 1973) considers systems of voting with lotteries as outcomes, and studies a property closely related to straightforwardness (1973). Gibbard (1973) discusses game forms with no element of chance, and proves that any straightforward game form of this kind is either dictatorial or duple. An equivalent theorem about non-manipulability of voting schemes is proved independently by Gibbard (1973) and by Satterthwaite (1975). The theorem by Gibbard (1976) on strategy-proof decision schemes has already been discussed here. The proof of the theorem in this paper follows the proof of the earlier theorem at a higher level of generality. Other work related to the topic of this paper is reviewed in that earlier paper (p. 6).

2. THEOREM AND PROOF

Let V be a finite non-empty set. A <u>lottery</u> ρ over V is a real-valued function whose domain is V, such that for each $x \in V$, $\rho(x) \geq 0$, and $\Sigma_{x \in V}\, \rho(x) = 1$. An <u>n-person profile set</u> $\mathcal{S}$ is a Cartesian product $S_1 \times \ldots \times S_n$, where each S_i is a finite non-empty set. A <u>game-form</u> is a function g such that for some positive integer n, the domain of g is an n-person profile set $\mathcal{S}$, and for some finite non-empty set V, the values of g are lotteries over V. n is called the <u>number of players</u> of g. V is called the <u>alternative set</u> of g, and members of V are called <u>alternatives</u> of g. Members of $\mathcal{S}$ are called <u>strategy-profiles</u> of g, and where $\mathcal{S} = S_1 \times \ldots \times S_n$, a member s_k of S_k is called a <u>strategy</u> for player k under g. Strategy profiles are writter in boldface (or on the typewriter with squiggly underlining) on the pattern $\mathbf{s} = \langle s_1, \ldots, s_n \rangle$. $s_k\mathbf{t}_{-k}$ is the strategy profile $\mathbf{u}$ such that $u_k = s_k$ and for all $i \neq k$, $u_i = t_i$. $\mathbf{t}_{-k}$ is called an <u>environment</u> for k. $g(x,\mathbf{s})$ is the probability $g(\mathbf{s})$ assigns to x.

A <u>utility scale</u> U over V is a real-valued function whose domain is V. Thus both utility scales over V and lotteries over V are vectors in the space $\mathbf{R}^V$ of all functions from V into the reals. Where U is a utility scale over V and ρ is a lottery over V, the <u>utility</u> of ρ <u>on scale</u> U is the inner product $U \cdot \rho = \Sigma_{x \in V}\, U(x)\rho(x)$. Let g be a game form with domain $S_1 \times \ldots \times S_n$ and alternative set V. Strategy $s_k \in S_k$ is U-<u>dominant</u> for k iff for every strategy profile $\underset{\sim}{t}$ of g, $U\cdot g(s_k\underset{\sim}{t}_{-k}) \geq U\cdot g(\underset{\sim}{t})$. g is <u>straightforward</u> iff for every k and U, there is a strategy $s_k \in S_k$ which is U-dominant for k. Game form g is <u>unilateral</u> iff there is a k such that for all s and t in its domain, if $s_k = t_k$ then $g(\underset{\sim}{s}) = g(\underset{\sim}{t})$. g is <u>duple</u> iff there are alternatives x and y such that for all $z \notin \{x,y\}$ and for all $\underset{\sim}{s}$ in its domain, $g(z,\underset{\sim}{s}) = 0$. g is a <u>probability mixture</u> of game forms $g_1, \ldots, g_\ell$ on domain $\mathcal{T}$ iff there is a lottery τ over $\{g_1, \ldots, g_\ell\}$ such that, where for each positive integer $\iota \leq \ell$, $\tau_\iota = \tau(g_\iota)$, for every $\underset{\sim}{s} \in \mathcal{T}$ and for every x, $g(x,\underset{\sim}{s}) = \tau_1 g_1(x,\underset{\sim}{s}) + \ldots + \tau_\ell g_\ell(x,\underset{\sim}{s})$. $D_k(s_k)$, the <u>domain of dominance</u> for k of strategy $s_k \in S_k$, is the set of all U such that s_k is U-dominant for k.

<u>Theorem</u>. Let g be a straightforward game form with domain $\mathcal{S} = S_1 \times \ldots \times S_n$ and alternative set V. For each player i, let T_i^* be the set of every $s_i \in S_i$ whose domain of dominance has a non-empty interior in $\underset{\sim}{R}^V$. Then (i) for any U and i, there is an $s_i \in T_i^*$ such that s_i is U-dominant for i, and (ii) on the domain $\mathcal{T}^* = T_1^* \times \ldots \times T_n^*$, g is a probability mixture of game forms each of which is either unilateral or duple.

<u>Notation</u>: Let m be $|V|$, the <u>number of alternatives</u>. Variables will range as follows:

x, y, z: Alternatives in V.

i, j, k: Positive integers $\leq n$, called <u>players</u>.

s_i, t_i, u_i, v_i: Strategies for i, which are members of S_i.

U: Utility scales over V.

Superscripts, primes, and the like do not affect the range of variables. Sections of the proof are numbered for reference.

1. Let $G_k(s_k, t_k, \underset{\sim}{v}_{-k})$, the <u>domain in which</u> s_k <u>is as good as</u> t_k in environment $\underset{\sim}{v}_{-k}$, be

$$\{U : U \cdot g(s_k \underset{\sim}{v}_{-k}) \geq U \cdot g(t_k \underset{\sim}{v}_{-k})\}.$$

Where $g(s_k \underset{\sim}{v}_{-k}) \neq g(t_k \underset{\sim}{v}_{-k})$, this is the closed half-space $\{U : U \cdot [g(s_k \underset{\sim}{v}_{-k}) - g(t_k \underset{\sim}{v}_{-k})] \geq 0\}$. Let $F_k(s_k \underset{\sim}{v}_{-k})$, the domain of maximality of s_k in environment $\underset{\sim}{v}_{-k}$ be $\bigcap \{G_k(s_k, t_k, v_{-k}) \mid t_k : t_k \in S_k\}$, the intersection of the domains in which s_k is as good as t_k in $\underset{\sim}{v}_{-k}$ for all t_k which are members of S_k. $D_k(s_k)$, the domain of dominance of s_k, is then $\bigcap \{F_k(s_k, \underset{\sim}{v}_{-k}) \mid \underset{\sim}{v}_{-k}\}$. $D_k(s_k)$ is thus a finite intersection of closed half-spaces; it is therefore closed and convex, and its boundaries are surfaces of dimension $m - 1$. To say that g is straightforward is to say that for every U and k, there is an $s_k \in S_k$ such that $U \in D_k(s_k)$, so that the sets $D_k(s_k)$ cover $\underset{\sim}{R}^V$.

2. The set T_k^* of members of S_k whose domains of dominance have interiors covers $\underset{\sim}{R}^V$. <u>Proof</u>: Let $H_1, \ldots, H_\ell$ be closed sets which cover $\underset{\sim}{R}^V$, and let sets $H_1, \ldots, H_p$ have interior points and $H_{p+1}, \ldots, H_\ell$ not. Then $H^* = H_{p+1} \cup \ldots \cup H_\ell$ has no interior points, and thus any point in H^* is in the closure of $\underset{\sim}{R}^V \setminus H^*$, which is a subset of $H_1 \cup \ldots \cup H_p$. Since these sets are closed, one of them contains each boundary point of their union. Since the sets $D_k(s_k)$ with $s_k \in S_k$ are closed and cover $\underset{\sim}{R}^V$, the assertion follows.

3. Strategies s_k and t_k are <u>equivalent</u> for k iff for every $\underset{\sim}{v}_{-k}$, $g(s_k \underset{\sim}{v}_{-k}) = g(t_k \underset{\sim}{v}_{-k})$; we write this $s_k \approx_k t_k$. Assertion: If $D_k(s_k)$ and $D_k(t_k)$ have an interior point in common, then $s_k \approx_k t_k$. <u>Proof</u>: If $g(s_k \underset{\sim}{v}_{-k}) \neq g(t_k \underset{\sim}{v}_{-k})$, then $G_k(s_k, t_k, \underset{\sim}{v}_{-k})$ and $G_k(t_k, s_k, \underset{\sim}{v}_{-k})$ have no interior points in common. Since $D_k(s_k) \subseteq G_k(s_k, t_k, \underset{\sim}{v}_{-k})$ and $D_k(t_k) \subseteq G_k(t_k, s_k, \underset{\sim}{v}_{-k})$, $D_k(s_k)$ and $D_k(t_k)$ have no interior points in common.

4. The relation $\approx_k$ divides T_k^* into equivalence classes. For each k, let T_k consist of one and only one member of each of these equivalence classes. Let $\mathcal{T} = T_1 \times \ldots \times T_n$. The sets $T_1, \ldots, T_n$ have these characteristics: $T_k \subseteq T_k^* \subseteq S_k$. $\underset{\sim}{R}^V$ is covered by sets $D_k(s_k)$ such that $s_k \in T_k$. For each $s_k \in T_k$, $D_k(s_k)$ has interior points. For any two members s_k, t_k of T_k, $D_k(s_k)$ and $D_k(t_k)$ have pairwise disjoint interiors.

5. Where s_k, $t_k \in T_k$ and $D_k(s_k)$ and $D_k(t_k)$ intersect in a surface of dimension $m - 1$, we say that s_k and t_k are <u>adjacent</u>. Let A_k be the set of all adjacent pairs of strategies in T_k. Where $\{s_k, t_k\} \in A_k$ or $s_k = t_k$, the <u>effect</u> $\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k})$ <u>of k's switching from</u> s_k <u>to</u> t_k <u>in</u> environment $\underset{\sim}{v}_{-k}$ is the vector $g(t_k \underset{\sim}{v}_{-k}) - g(s_k \underset{\sim}{v}_{-k})$. We now show that for adjacent strategies, the direction of the effect of a switch is independent of environment.

6. Let $\{s_k, t_k\} \in A_k$, and let $\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k}) \neq 0$. Then $\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k})$ is orthogonal to the boundary $D_k(s_k) \cap D_k(t_k)$, and points in the sense from $D_k(s_k)$ to $D_k(t_k)$. <u>Proof</u>: Since $D_k(s_k) \subseteq G_k(s_k, t_k, \underset{\sim}{v}_{-k})$ and $D_k(t_k) \subseteq G_k(t_k, s_k, \underset{\sim}{v}_{-k})$, where $E = G_k(s_k, t_k, \underset{\sim}{v}_{-k}) \cap G_k(t_k, s_k, \underset{\sim}{v}_{-k})$, we have $D_k(s_k) \cap D_k(t_k) \subseteq E$. Now E is the set

$$\{U : U \cdot g(s_k \underset{\sim}{v}_{-k}) = U \cdot g(t_k \underset{\sim}{v}_{-k})\};$$

thus for $U \in E$, $U \cdot [g(t_k \underset{\sim}{v}_{-k}) - g(s_k \underset{\sim}{v}_{-k})] = 0$, and E is orthogonal to $\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k})$. Let U_t lie in the interior of $D_k(t_k)$ and U_s lie in the interior of $D_k(s_k)$, with $U_t - U_s$ orthogonal to E. Then $U_t \cdot g(t_k \underset{\sim}{v}_{-k}) > U_t \cdot g(s_k \underset{\sim}{v}_{-k})$ and $U_s \cdot g(t_k \underset{\sim}{v}_{-k}) < U_s \cdot g(s_k \underset{\sim}{v}_{-k})$. Therefore

$$(U_t - U_s) \cdot [g(t_k \underset{\sim}{v}_{-k}) - g(s_k \underset{\sim}{v}_{-k})] > 0,$$

and $\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k})$ has the sense of $U_t - U_s$.

7. Any lottery lies in the simplex C^{m-1} of $\rho \in \underset{\sim}{R}^{V}$ such that $\Sigma_x \rho(x) = 1$ and $\rho(x) \geq 0$ for each x. This simplex is of dimension $m-1$. A <u>direction</u> φ of C^{m-1} is a vector of length one such that $\Sigma_x \varphi(x) = 0$. We have shown that for each pair $\{s_k, t_k\} \in A_k$, there is a direction $\theta_k(s_k, t_k)$ which has this property: for every $\underset{\sim}{v}_{-k}$, there is a real $\alpha > 0$ such that $\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k}) = \alpha\theta_k(s_k, t_k)$.

8. Now for a fixed environment $\underset{\sim}{v}_{-k}$, consider the set of lotteries $g(s_k \underset{\sim}{v}_{-k})$ that k can attain with various strategies $s_k \in T_k$. They form the vertices of a convex polyhedron $H_k(\underset{\sim}{v}_{-k})$ in C^{m-1} (where more than one strategy may coincide at a vertex). For if $g(s_k \underset{\sim}{v}_{-k})$ is a convex combination of other distinct $g(t_k \underset{\sim}{v}_{-k})$ and $g(u_k \underset{\sim}{v}_{-k})$, it can be maximal only on $G_k(t_k, u_k, \underset{\sim}{v}_{-k}) \cap G_k(u_k, t_k, \underset{\sim}{v}_{-k})$, and hence its domain of maximality has no interior. A <u>course</u> from vertex p to vertex q of a polyhedron H is a sequence $p_0, \ldots, p_\nu$ such that $p_0 = p$, $p_\nu = q$, and each line segment $p_{\iota-1}p_\iota$ is an edge of H. A line segment pq lies in direction φ iff for some real α, $p - q = \alpha\varphi$. The following lemma about convex polyhedra is proved in the Appendix to the proof.

<u>Lemma</u> 1. Let H be a convex polyhedron, and where ψ and φ are directions, let p and q be vertices such that p is the sole ψ-maximal vertex of H and q is a φ-minimal vertex of H. Then there is a course $p_0, \ldots, p_\nu$ on H from q to p such that (i) for each ι such that $1 \leq \iota \leq \nu$, $\varphi \cdot (p_\iota - p_{\iota-1}) > 0$, and (ii) if $\psi \cdot \varphi \leq 0$, then none of the edges $p_{\iota-1}p_\iota$ lie in direction φ.

9. For any k, s_k, and φ, either (i) for every v_{-k}, $g(s_k v_{-k})$ is connected to every φ-minimal vertex of $H_k(v_{-k})$ by a course none of whose segments lie in direction φ, or (ii) for every v_{-k}, $g(s_k v_{-k})$ is connected to every φ-maximal vertex of $H_k(v_{-k})$ by a course none of whose segments lie in direction φ. Proof: Since $D_k(s_k)$ has interior points, there is a direction ψ in the interior of $D_k(s_k)$ such that $\psi \cdot \varphi \neq 0$. For any v_{-k}, $g(s_k v_{-k})$ is a uniquely ψ-maximal vertex of $H_k(v_{-k})$. By Lemma 1, if $\psi \cdot \varphi < 0$, then $g(s_k v_{-k})$ is connected to every φ-minimal vertex of $H_k(v_{-k})$ by a course none of whose segments lie in direction φ. If $\varphi \cdot \psi > 0$, then $\psi \cdot (-\varphi) < 0$, and $g(s_k v_{-k})$ is connected to every $(-\varphi)$-minimal vertex by such a course.

10. For each i and φ, we now let $T_i^{-\varphi}$ be the set of $s_k \in T_k$ which satisfy (i) of 9, and let $T_i^{+\varphi} = T_i \setminus T_i^{-\varphi}$. $T_i^{+\varphi}$ may be empty; if it has members, then they satisfy (ii) of 9. It follows that for any two members s_k and t_k of $T_k^{-\varphi}$ and for any v_{-k}, there is a course from $g(s_k v_{-k})$ to $g(t_k v_{-k})$ on $H_k(v_{-k})$ none of whose segments lie in direction φ. The same holds for any two members of $T_i^{+\varphi}$. If $T_i^{+\varphi}$ is non-empty, we say that T_i is φ-<u>separable</u>. Clearly T_i is $(-\varphi)$-separable iff T_i is φ-separable.

11. Where $\varphi = \theta_k(s_k, t_k)$, let $v_i \in T_i^{-\varphi}$ iff $w_i \in T_i^{-\varphi}$ for all $i \neq k$. Then $\epsilon_k(s_k, t_k, v_{-k}) = \epsilon_k(s_k, t_k, w_{-k})$.
Proof: We first prove the assertion for the case where v_{-k} and w_{-k} differ only by a single person j's switching v_j to w_j, where $j \neq k$, $\{v_j, w_j\} \in A_j$, and $\theta_j(v_j, w_j) \neq \pm\theta_k(s_k, t_k)$.

Consider four lotteries: $p_1 = g(s_j s_k \underset{\sim}{v}_{-j-k})$, $p_2 = g(s_j t_k \underset{\sim}{v}_{-j-k})$, $p_3 = g(t_j t_k \underset{\sim}{v}_{-j-k})$, and $p_4 = g(s_j t_k \underset{\sim}{v}_{-j-k})$. $p_1 p_2 p_3 p_4$ form a quadralateral which, by $\underset{\sim}{6}$, has opposite sides parallel. Hence opposite sides are equal: j's switch does not alter the effect of k's switch.

Now let $v_j \in T_j$ iff $w_j \in T_j$. Then w_j can be obtained from from s_j by a sequence of switches, none of which is in direction $\pm\varphi$. Hence none alters the effect of k's switch in direction φ. The same argument applies to every other $i \neq k$, and the assertion of this section is proved.

$\underset{\sim}{12}$. For the remainder of the proof we adopt these notations. Let $\underset{\sim}{v}^*$ be an arbitrary fixed strategy profile. If $T_i^{+\varphi}$ is empty, let $u_i^{-\varphi}$ be an arbitrary fixed member of T_i. If $T_i^{+\varphi}$ is non-empty, then let $u_i^{-\varphi}$ and $u_i^{+\varphi}$ be such that for all pairs $\langle s_i, t_i \rangle$ where $\{s_i, t_i\} \in A_i$ and and $\theta_i(s_i, t_i) = \varphi$, $|g(t_i \underset{\sim}{v}^*_{-i}) - g(s_i \underset{\sim}{v}^*_{-i})|$ is minimal when $s_i = u_i^{-\varphi}$ and $t_i = u_i^{+\varphi}$. We stipulate further that if $\psi = -\varphi$, then $u_i^{+\psi} = u_i^{-\varphi}$ and $u_i^{+\psi} = u_i^{-\varphi}$. An <u>alignment</u> ω is a pair of directions $\{\varphi, -\varphi\}$. The variable ω will range over alignments such that for some k, s_k, and t_k with $\{s_k, t_k\} \in A_k$, $\theta_k(s_k, t_k) \in \omega$. Clearly the class of such alignments is finite. $\mathcal{T}$ is <u>ω-separable</u> iff for some i and some $\varphi \in \omega$, T_i is φ-separable. For any s_i and ω with $\varphi \in \omega$, let $s_i^{\omega} = u_i^{-\varphi}$ iff $s_i \in T_i^{-\varphi}$ and let $s_i = u_i^{+\varphi}$ iff $s_i \in T_i^{+\varphi}$.

13. If $\{s_k, t_k\} \in A_k$ and $\theta_k(s_k, t_k) = \varphi \in \omega$, then for any two environments $\underset{\sim}{v}_{-k}$ and $\underset{\sim}{w}_{-k}$,

$$\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k}) - \epsilon_k(s_k^{\omega}, t_k^{\omega}, \underset{\sim}{v}_{-k}^{\omega}) = \epsilon_k(s_k, t_k, \underset{\sim}{w}_{-k}) - \epsilon_k(s_k^{\omega}, t_k^{\omega}, \underset{\sim}{w}_{-k}^{\omega}).$$

Proof: We know from 11 that $\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k}^{\omega}) = \epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k})$ and $\epsilon_k(s_k, t_k, \underset{\sim}{w}_{-k}^{\omega}) = \epsilon_k(s_k, t_k, \underset{\sim}{w}_{-k})$. We need only show, then, that

$$\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k}^{\omega}) - \epsilon_k(s_k^{\omega}, t_k^{\omega}, \underset{\sim}{v}_{-k}^{\omega}) = \epsilon_k(s_k, t_k, \underset{\sim}{w}_{-k}^{\omega}) - \epsilon_k(s_k^{\omega}, t_k^{\omega}, \underset{\sim}{w}_{-k}^{\omega}). \quad (1)$$

We have

$$\epsilon_k(s_k, t_k, \underset{\sim}{v}_{-k}^{\omega}) = g(t_k \underset{\sim}{v}_{-k}^{\omega}) - g(s_k \underset{\sim}{v}_{-k}^{\omega});$$

$$\epsilon_k(s_k, t_k, \underset{\sim}{w}_{-k}^{\omega}) = g(t_k \underset{\sim}{w}_{-k}^{\omega}) - g(s_k \underset{\sim}{w}_{-k}^{\omega}).$$

Go now from s_k to s_k^{ω} by switches in directions other than $\pm\varphi$, from s_k^{ω} to t_k^{ω}, and from t_k^{ω} to t_k by switches in directions other than $\pm\varphi$. $\underset{\sim}{v}_{-k}^{\omega}$ and $\underset{\sim}{w}_{-k}^{\omega}$ differ from each other only by switches in the directions $\pm\varphi$; thus the switches in the sequence from s_k to s_k^{ω} have equal effects in both environments. Therefore

$$g(s_k \underset{\sim}{v}_{-k}^{\omega}) - g(s_k^{\omega} \underset{\sim}{v}_{-k}^{\omega}) = g(s_k \underset{\sim}{w}_{-k}^{\omega}) - g(s_k^{\omega} \underset{\sim}{w}_{-k}^{\omega}). \quad (2)$$

For the same reason,

$$g(t_k \underset{\sim}{v}_{-k}^{\omega}) - g(t_k^{\omega} \underset{\sim}{v}_{-k}^{\omega}) = g(t_k \underset{\sim}{w}_{-k}^{\omega}) - g(t_k^{\omega} \underset{\sim}{w}_{-k}^{\omega}). \quad (3)$$

Adding (2) and (3), we get

$$[g(t_k\underset{\sim}{v}^{\omega}_{-k}) - g(s_k\underset{\sim}{v}^{\omega}_{-k})] - [g(t^{\omega}_k\underset{\sim}{v}^{\omega}_{-k}) - g(s^{\omega}_k\underset{\sim}{v}^{\omega}_{-k})] =$$

$$[g(t_k\underset{\sim}{w}^{\omega}_{-k}) - g(s_k\underset{\sim}{w}^{\omega}_{-k})] - [g(t^{\omega}_k\underset{\sim}{w}^{\omega}_{-k}) - g(s^{\omega}_k\underset{\sim}{w}^{\omega}_{-k})],$$

which is (1).

14. Thus where $\theta_k(s_k,t_k) \in \omega$, $\epsilon_k(s_k,t_k,\underset{\sim}{v}_{-k}) - \epsilon_k(s^{\omega}_k,t^{\omega}_k,\underset{\sim}{v}^{\omega}_{-k})$ is a constant which is independent of the environment $\underset{\sim}{v}_{-k}$: it depends only on s_k and t_k. Call this function $\delta_k(s_k,t_k)$; it is defined for all adjacent pairs of strategies for k. On the other hand, the function $\epsilon_k(s^{\omega}_k,t^{\omega}_k,\underset{\sim}{v}^{\omega}_{-k})$ depends only on whether s_k, t_k, and the various v_i's for $i \neq k$ are in $T_i^{-\varphi}$ or $T_i^{+\varphi}$ sets. Call this function $\gamma_k(s_k,t_k,\underset{\sim}{v}_{-k})$; it is defined, then, as $\epsilon_k(s^{\omega}_k,t^{\omega}_k,\underset{\sim}{v}^{\omega}_{-k})$ where $\theta_k(s_k,t_k) \in \omega$. Since for the $\underset{\sim}{v}^*$ introduced earlier, $u_k^{-\varphi}$ and $u_k^{+\varphi}$, if both are defined, were chosen so as to make $|\epsilon_k(u_k^{-\varphi},u_k^{+\varphi},\underset{\sim}{v}^*_{-k})|$ minimal for $\theta_k(u_k^{-\varphi},u_k^{+\varphi}) = \varphi$, whenever $\theta_k(s_k,t_k) = \varphi$, we have $|\epsilon_k(s_k,t_k,\underset{\sim}{v}^*_{-k})| \geq |\epsilon_k(u_k^{-\varphi},u_k^{+\varphi},\underset{\sim}{v}^*_{-k})|$. Thus if $t_k \in T_k^{+\varphi}$, then $|\epsilon_k(s_k,t_k,\underset{\sim}{v}_{-k})| \geq |\epsilon_k(s^{\omega}_k,t^{\omega}_k,\underset{\sim}{v}^{\omega}_{-k})|$, and since both effects are in the $\theta_k(s_k,t_k)$ direction and their difference is $\delta_k(s_k,t_k)$, it follows that $\delta_k(s_k,t_k)$, if non-zero, is a vector in the $\theta_k(s_k,t_k)$ direction. In short, then, both $\delta_k(s_k,t_k)$ and $\gamma_k(s_k,t_k,\underset{\sim}{v}_{-k})$ are vectors which, if non-zero, are in the $\theta_k(s_k,t_k)$ direction, and their sum is $\epsilon_k(s_k,t_k,\underset{\sim}{v}_{-k})$.

15. Let a <u>path</u> from s_k to t_k be a sequence $s_k^0, \ldots, s_k^{\lambda}$ of strategies in T_k such that $s_k^0 = s_k$, $s_k^{\lambda} = t_k$, and $\{s_k^{\iota-1}, s_k^{\iota}\} \in A_k$ for each ι with $0 \leq \iota \leq \lambda$.

A <u>circuit</u> is a path from a strategy to itself. For any circuit $s_k^0, \ldots, s_k^\lambda$ from s_k^0 to itself,

$$\Sigma_{\iota=1}^{\lambda} \delta_k(s_k^{\iota-1}, s_k^\iota) = 0. \tag{4}$$

<u>Proof</u>: We have at each step in the circuit

$$\delta_k(s_k^{\iota-1}, s_k^\iota) = \epsilon_k(s_k^{\iota-1}, s_k^\iota, \underset{\sim}{v}_{-k}) - \gamma_k(s_k^{\iota-1}, s_k^\iota, \underset{\sim}{v}_{-k}).$$

But where $\varphi(\iota) = \theta_k(s_k^{\iota-1}, s_k^\iota) \in \omega(\iota)$, we have

$$\gamma_k(s_k^{\iota-1}, s_k^\iota, \underset{\sim}{v}_{-k}) = g(s_k^{\iota\omega(\iota)} \underset{\sim}{v}_{-k}^{\omega(\iota)}) - g(s_k^{(\iota-1)\omega(\iota)} \underset{\sim}{v}_{-k}^{\omega(\iota)}).$$

Therefore

$$\Sigma_{\iota=1}^{\lambda} \gamma_k(s_k^{\iota-1}, s_k^\iota, \underset{\sim}{v}_{-k}) = \Sigma_{\iota=1}^{\lambda}[g(s_k^{\iota\omega(\iota)} \underset{\sim}{v}_{-k}^{\omega(\iota)}) - g(s_k^{(\iota-1)\omega(\iota)} \underset{\sim}{v}_{-k}^{\omega(\iota)})].$$

$$= \Sigma_\omega \Sigma\{g(s_k^{\iota\omega} \underset{\sim}{v}_{-k}^{\omega}) - g(s_k^{(\iota-1)\omega} \underset{\sim}{v}_{-k}^{\omega}) \mid \iota : \varphi(\iota) \in \omega\}.$$

But for $\varphi(\iota) \notin \omega$, we have $s_k^{(\iota-1)\omega} = s_k^\omega$. For since $s_k^{\iota-1}$ and s_k^ι are adjacent and $\theta_k(s_k^{\iota-1}, s_k^\iota) \in \omega$, where $\varphi \in \omega$, either $s_k^{\iota-1}$ and s_k^ι are both in $T_i^{-\varphi}$, or $s_k^{\iota-1}$ and s_k^ι are both in $T_i^{+\varphi}$. In the first case, $s_k^{\iota-1} = u_i^{-\varphi} = s_k^\iota$, and in the second, $s_k^{\iota-1} = u_i^{+\varphi} = s_k^\iota$. Thus

$$\Sigma\{g(s_k^{\iota\omega} \underset{\sim}{v}_{-k}^{\omega}) - g(s_k^{(\iota-1)\omega} \underset{\sim}{v}_{-k}^{\omega}) \mid \iota : \varphi(\iota) \in \omega\}$$

$$= \Sigma_{\iota=1}^{\lambda} [g(s_k^{\lambda\omega} \underset{\sim}{v}_{-k}^{\omega}) - g(s_k^{(\iota-1)\omega} \underset{\sim}{v}_{-k}^{\omega})]$$

$$= g(s_k^{\lambda\omega} \underset{\sim}{v}_{-k}^{\omega}) - g(s_k^{0\omega} \underset{\sim}{v}_{-k}^{\omega}) = 0.$$

Therefore $\Sigma_{\iota=1}^{\lambda} \gamma_k(s_k^{\iota-1}, s_k^\iota, \underset{\sim}{v}_{-k}) = 0$. Since also, $\Sigma_{\iota=1}^{\lambda} \epsilon_k(s_k^{\iota-1}, s_k^\iota, \underset{\sim}{v}_{-k}) = 0$, (4) holds.

16. Since a circuit of δ_k's sums to zero, and from the way δ_k is defined, $\delta_k(s_k,t_k) = -\delta_k(t_k,s_k)$, it follows that a sum of δ_k's is <u>path independent</u>, in the sense that for any two paths $s_k^0,\ldots,s_k^\lambda$ and $t_k^0,\ldots,t_k^\lambda$ from s_k to t_k,

$$\Sigma_{\iota=1}^{\lambda}\ \delta_k(s_k^{\iota-1},s_k^{\iota}) = \Sigma_{\iota=1}^{\lambda}\ \delta_k(t_k^{\iota-1},t_k^{\iota}).$$

Call this quantity $f_k^*(t_k,s_k)$. For any orientation ω with $\mathcal{T}$ ω-separable, let $h_\omega^*(\underset{\sim}{t},\underset{\sim}{s}) = g(\underset{\sim}{t}^{\omega}) - g(\underset{\sim}{s}^{\omega})$.

17. Differences in values of g decompose in this way:

$$g(\underset{\sim}{t}) - g(\underset{\sim}{s}) = \Sigma_k\ f_k^*(t_k,s_k) + \Sigma_\omega h_\omega^*(\underset{\sim}{t},\underset{\sim}{s}).$$

<u>Proof</u>: Let $\underset{\sim}{s}^0,\ldots,\underset{\sim}{s}^\lambda$ be a sequence with $\underset{\sim}{s}^0 = \underset{\sim}{s}$, $\underset{\sim}{s}^\lambda = \underset{\sim}{t}$, and with $\underset{\sim}{s}^\iota$ differing from $\underset{\sim}{s}^{\iota-1}$ only in that $k(\iota)$ switches in a direction $\varphi\in\omega(\iota)$. We have

$$\begin{aligned} g(\underset{\sim}{s}^\iota) - g(\underset{\sim}{s}^{\iota-1}) &= \epsilon_{k(\iota)}(s_{k(\iota)}^{\iota-1},s_{k(\iota)}^{\iota},\underset{\sim}{s}_{-k(\iota)}^{\iota}) \\ &= \delta_{k(\iota)}(s_{k(\iota)}^{\iota-1},s_{k(\iota)}^{\iota}) + \gamma_{k(\iota)}(s_{k(\iota)}^{\iota-1},s_{k(\iota)}^{\iota},\underset{\sim}{s}_{-k(\iota)}^{\iota}) \\ &= [f_{k(\iota)}^*(\underset{\sim}{s}^\iota,\underset{\sim}{s}) - f_{k(\iota)}^*(\underset{\sim}{s}^{\iota-1},\underset{\sim}{s})] + [h_{\omega(\iota)}(\underset{\sim}{s}^\iota,\underset{\sim}{s}) - h_{\omega(\iota)}(\underset{\sim}{s}^{\iota-1},\underset{\sim}{s})] \end{aligned}$$

For $i\neq k(\iota)$, $f_i^*(\underset{\sim}{s}^\iota,\underset{\sim}{s}) = f_i^*(\underset{\sim}{s}^{\iota-1},\underset{\sim}{s})$, and for $\omega\neq\omega(\iota)$, $h_\omega(\underset{\sim}{s}^\iota,\underset{\sim}{s}) = h_\omega(\underset{\sim}{s}^{\iota-1},\underset{\sim}{s})$. Therefore $g(\underset{\sim}{s}^\iota) - g(\underset{\sim}{s}^{\iota-1})$ equals

$$\Sigma_i f_i^*(\underset{\sim}{s}^\iota,\underset{\sim}{s}) - \Sigma_i f_i^*(\underset{\sim}{s}^{\iota-1},\underset{\sim}{s}) + \Sigma_\omega h_\omega^*(\underset{\sim}{s}^\iota,\underset{\sim}{s}) - \Sigma_\omega h_\omega^*(\underset{\sim}{s}^{\iota-1},\underset{\sim}{s}),$$

and $g(\underset{\sim}{s}^\iota) - g(\underset{\sim}{s})$ equals

$$\Sigma_i f_i^*(\underset{\sim}{s}^\iota,\underset{\sim}{s}) - \Sigma_i f_i^*(\underset{\sim}{s}^0,\underset{\sim}{s}) + \Sigma_\omega h_\omega^*(\underset{\sim}{s}^\iota,\underset{\sim}{s}) - \Sigma_\omega h_\omega^*(\underset{\sim}{s}^0,\underset{\sim}{s})$$

$$= \Sigma_i f_i^*(\underset{\sim}{s}^\iota,\underset{\sim}{s}) + \Sigma_\omega h_\omega^*(\underset{\sim}{s}^\iota,\underset{\sim}{s}).$$

18. If for fixed x, $g(x,\underset{\sim}{s})$ is minimal for $\underset{\sim}{s} = \underset{\sim}{s}^x$, then where x and $\underset{\sim}{t}$ are fixed, for each k, $f_k^*(x,\underset{\sim}{s},\underset{\sim}{t})$ is minimal for $\underset{\sim}{s} = \underset{\sim}{s}^x$, and for each ω, $h_\omega^*(x,\underset{\sim}{s},\underset{\sim}{t})$ is minimal for $\underset{\sim}{s} = \underset{\sim}{s}^x$. Proof: Where $\varphi \in \omega$, all values of $h_\omega^*(\underset{\sim}{s},\underset{\sim}{t})$ are of the form $\alpha\varphi$, where α is a real number. If $\varphi(x) > 0$, then for fixed $\underset{\sim}{t}$, $h_\omega^*(x,\underset{\sim}{s},\underset{\sim}{t})$ is minimal when $s_i \in T_i^{-\varphi}$ for all i. Where η_x is the unit vector in the x direction, $\eta_x \cdot \varphi > 0$, and hence by Lemma 1 and the definition of $T_i^{-\varphi}$, $s_i^x \in T_i^{-\varphi}$. Therefore $h_\omega^*(x,\underset{\sim}{s},\underset{\sim}{t})$ is minimal for $\underset{\sim}{s} = \underset{\sim}{s}^x$. If $\varphi(x) < 0$, repeat the argument using $\psi = -\varphi$. If $\varphi(x) = 0$, then $h_\omega^*(x,\underset{\sim}{s},\underset{\sim}{t}) = \alpha\varphi(x) = 0$, and so any value of s minimizes it.

Now go from s_k^x to s_k in a path which makes $\eta_x \cdot g(s_k;\underset{\sim}{v}^*_{-k})$ non-decreasing; the existence of such a path is guaranteed by (i) of Lemma 1. Since switches in $f_k^*(s_k,t_k)$ are always in the same direction as switches in g (by $\underset{\sim}{14}$), these switches are all non-x-decreasing, and $f_k^*(x,s_k,t_k) \geq f_k^*(x,s_k^x,t_k)$.

19. A <u>measure</u> over V is a function τ whose domain is V and whose values are non-negative real numbers; the <u>weight</u> of a measure τ is $\Sigma_x \tau(x)$. A <u>partial</u> <u>game</u> <u>form</u> on $\mathcal{T} = T_1 \times \ldots \times T_n$ and V is a function whose domain is $\mathcal{T}$ and whose values, for some $\alpha \geq 0$, are measures over V of weight α. Let f be a partial game form on $\mathcal{T}$ and V. f is <u>alternative-eliminating</u> iff for each $x \in V$ there is an $\underset{\sim}{s} \in \mathcal{T}$ such that $f(x,\underset{\sim}{s}) = 0$. f is <u>linear</u> iff there

are a real-valued function α with $0 \leq \alpha(s) \leq 1$ and strategy profiles $\underset{\sim}{t}$ and $\underset{\sim}{t}^* \in \mathcal{T}$ such that for all $\underset{\sim}{s} \in \mathcal{T}$, $f(\underset{\sim}{s}) = [1 - \alpha(\underset{\sim}{s})]f(\underset{\sim}{t}) + \alpha(\underset{\sim}{s})f(\underset{\sim}{t}^*)$.

20. Now for all $\underset{\sim}{s}$ and x, let $h_\omega(x,\underset{\sim}{s}) = h^*_\omega(x,\underset{\sim}{s},\underset{\sim}{s}^x)$. Then by $\underset{\sim}{18}$, $h_\omega(x,\underset{\sim}{s}) \geq 0$ for all x and $\underset{\sim}{s}$. Moreover, from the way h^*_ω is defined, $h_\omega(x,\underset{\sim}{s}) = g(x,\underset{\sim}{s}^\omega) - g(x,\underset{\sim}{s}^{x\omega})$, and

$$\Sigma_x h_\omega(x,\underset{\sim}{s}) = \Sigma_x[g(x,\underset{\sim}{s}^\omega) - g(x,\underset{\sim}{s}^{x\omega})] = 1 - \Sigma_x g(x,\underset{\sim}{s}^{x\omega}),$$

which is independent of ω. Thus h_ω is a partial game form. Since for all x, $h_\omega(x,\underset{\sim}{s}^x) = h^*_\omega(x,\underset{\sim}{s}^x,\underset{\sim}{s}^x) = 0$, h_ω is alternative-eliminating. h_ω is linear, since for any $\underset{\sim}{s}$ and $\underset{\sim}{t}$, $h_\omega(\underset{\sim}{s}) = h^*_\omega(\underset{\sim}{s},\underset{\sim}{t}) + h_\omega(\underset{\sim}{t})$, $h^*_\omega(\underset{\sim}{s},\underset{\sim}{t}) = g(\underset{\sim}{t}^\omega) - g(\underset{\sim}{s}^\omega)$, and for any i, s_i, and t_i with $\{s_i,t_i\} \in A_i$, either $s_i = t_i$ or $\theta_i(s_i,t_i) \in \omega$.

21. Similarly, let $f_k(x,\underset{\sim}{s}) = f^*_k(x,\underset{\sim}{s},\underset{\sim}{s}^x)$ for each x. From $\underset{\sim}{18}$, it follows that $f_k(x,\underset{\sim}{s}) \geq 0$ for all x and $\underset{\sim}{s}$. By path-independence,

$$f^*_k(x,\underset{\sim}{s},\underset{\sim}{s}^x) = f^*_k(x,\underset{\sim}{s}^x,\underset{\sim}{v}^*) + f^*_k(x,\underset{\sim}{v}^*,\underset{\sim}{s}^x).$$

Hence $\Sigma_x f_k(x,\underset{\sim}{s}) = \Sigma_x f^*_k(x,\underset{\sim}{s},\underset{\sim}{v}^*) + \Sigma_x f^*_k(x,\underset{\sim}{v}^*,\underset{\sim}{s}^x)$. Since $\Sigma_x f^*_k(x,\underset{\sim}{s},\underset{\sim}{v}^*) = 0$, this is $\Sigma_x f^*_k(x,\underset{\sim}{v}^*,\underset{\sim}{s}^x)$, which is independent of $\underset{\sim}{s}$. Thus each f_k is a partial game form. f_k is unilateral, since it is characterized in terms of f^*_k, which in turn is characterized in terms of δ_k, whose values are environment-independent.

22. Let $f_0(x,\underset{\sim}{s}) = g(x,\underset{\sim}{s}^x)$ for all x. Then $f_0(x,\underset{\sim}{s}) \geq 0$ for all x and $\underset{\sim}{s}$, and $\Sigma_x f_0(x,\underset{\sim}{s})$ is independent of $\underset{\sim}{s}$. Thus f_0 is a partial game form. Note also that f_0 is unilateral in a degenerate way. Now since

$$\begin{aligned} g(x,\underset{\sim}{s}) - g(x,\underset{\sim}{s}^x) &= \Sigma_k f_k^*(x,\underset{\sim}{s},\underset{\sim}{s}^x) + \Sigma_\omega h_\omega^*(x,\underset{\sim}{s},\underset{\sim}{s}^x) \\ &= \Sigma_k f_k(x,\underset{\sim}{s}) + \Sigma_\omega h_\omega(x,\underset{\sim}{s}), \end{aligned}$$

and $g(x,\underset{\sim}{s}^x) = f_0(x,\underset{\sim}{s})$, we have

$$g(x,\underset{\sim}{s}) = f_0(x,\underset{\sim}{s}) + \Sigma_k f_k(x,\underset{\sim}{s}) + \Sigma_\omega h_\omega(x,\underset{\sim}{s}).$$

Thus on $\mathcal{T}$, g is a finite sum of partial game forms, each of which is either unilateral or both linear and alternative-eliminating.

23. Any linear alternative-eliminating partial game form is a sum of duple partial game forms. <u>Proof</u>: Let ρ^- and ρ^+ be the extreme values of $h_\omega(\underset{\sim}{s})$. Since h_ω is linear, there is a real-valued function $\alpha(\underset{\sim}{s})$ such that for all $\underset{\sim}{s}$, $0 \leq \alpha(\underset{\sim}{s}) \leq 1$ and $h_\omega(\underset{\sim}{s}) = [1-\alpha(s)]\rho^- + \alpha(s)\rho^+$, and since in addition h_ω is alternative-eliminating, for each x, either $\rho^-(x) = 0$ or $\rho^+(x) = 0$.

Where the alternatives in V are $x_1, \ldots, x_m$, let $\sigma_\kappa^+ = \Sigma_{\iota=1}^\kappa \rho^+(x_\iota)$ and let $\sigma_\lambda^- = \Sigma_{\iota=1}^\lambda \rho^-(x_\iota)$; let $\sigma_0^+ = 0 = \sigma_0^-$. Let $\sigma_1, \ldots, \sigma_l$ be the values of σ_κ^+ and σ_λ^- in order of magnitude. Then for each κ, the interval $(\sigma_{\kappa-1}^+, \sigma_\kappa^+]$ consists of zero or more intervals of the form $(\sigma_{\iota-1}, \sigma_\iota]$. Where $(\sigma_{\iota-1}, \sigma_\iota] \subseteq (\sigma_{\kappa-1}^+, \sigma_\kappa^+]$, we shall say $\kappa(\iota) = \kappa$.

We have

$$\Sigma\{\sigma_\iota - \sigma_{\iota-1} \mid \iota : \kappa(\iota) = \kappa\} = \sigma^+_\kappa - \sigma^+_{\kappa-1} = \rho^+(x_\kappa). \tag{5}$$

Likewise for each λ, the interval $(\sigma^-_{\lambda-1}, \sigma^-_\lambda]$ consists of zero or more intervals of the form $(\sigma_{\iota-1}, \sigma_\iota)$. Let $\lambda(\iota) = \lambda$ iff $(\sigma_{\iota-1}, \sigma_\iota] \subseteq (\sigma^-_{\iota-1}, \sigma^-_\iota]$; then

$$\Sigma\{\sigma - \sigma_{\iota-1} \mid \iota : \lambda(\iota) = \lambda\} = \rho^-(x_\lambda). \tag{6}$$

For each ι: $0 \leq \iota \leq l$, define a duple partial game form d_ι as follows.

$$d_\iota(x_{\kappa(\iota)}, \underset{\sim}{s}) = \alpha(\underset{\sim}{s})(\sigma_\iota - \sigma_{\iota-1}).$$

$$d_\iota(x_{\lambda(\iota)}, \underset{\sim}{s}) = [1 - \alpha(s)](\sigma_\iota - \sigma_{\iota-1}).$$

$$d_\iota(y, \underset{\sim}{s}) = 0 \quad \text{for} \quad y \notin \{x_{\kappa(\iota)}, x_{\lambda(\iota)}\}.$$

This is indeed a partial game form. Its values are non-negative, since $\alpha(\underset{\sim}{s})$, $1 - \alpha(\underset{\sim}{s})$, and $\sigma_\iota - \sigma_{\iota-1}$ are always non-negative. The sum of its values is $\sigma_\iota - \sigma_{\iota-1}$, which is constant for all $\underset{\sim}{s}$. Now for any x_κ, $\Sigma_{\iota=1} d_\iota(x_\kappa, \underset{\sim}{s})$ is

$$\alpha(\underset{\sim}{s})\Sigma\{\sigma_\iota - \sigma_{\iota-1} \mid \iota : \kappa = \kappa(\iota)\} + [1 - \alpha(\underset{\sim}{s})]\Sigma\{\sigma_\iota - \sigma_{\iota-1} \mid \iota : \kappa = \lambda(\iota)\}$$

$$= \alpha(\underset{\sim}{s})\rho^+(x_\kappa) + [1 - \alpha(\underset{\sim}{s})]\rho^-(x_\kappa),$$

with the last step following from (5) and (6). Thus

$$\Sigma^l_{\iota=1} d_\iota(\underset{\sim}{s}) = \alpha(\underset{\sim}{s})\rho^+ + [1 - \alpha(\underset{\sim}{s})]\rho = h_\omega(\underset{\sim}{s}).$$

24. Proof of Theorem: We have seen that on $\mathcal{T}$, g is a sum of partial forms, each of which is either unilateral or duple. Partial forms of weight zero do not contribute to this sum. Let $g = c_1' + \ldots + c_l'$, where these partial forms have positive weights $\alpha_1, \ldots, \alpha_l$. Then let $c_\iota = (1/\alpha_\iota)c_\iota$ for each $\iota : 1 \leq \iota \leq l$; then

$$g = \alpha_1 c_1 + \ldots + \alpha_l c_l, \tag{7}$$

where each c_ι is unilateral or duple. We may expand the domain of each c_ι to $\mathcal{T}^*$ by treating equivalent strategies alike: if $\underset{\sim}{s} \in \mathcal{T}^*$, take the $\underset{\sim}{s}' \in \mathcal{T}$ such that $s_1' \approx_1 s_1, \ldots, s_n' \approx_n s_n$, and let $c_\iota(\underset{\sim}{s}) = c_\iota(\underset{\sim}{s}')$ for each ι. Then since $g(\underset{\sim}{s}) = g(\underset{\sim}{s}')$, (7) still holds. That proves the Theorem.

Appendix to the proof: Lemma 1. Let H be a convex polyhedron, and where ψ and φ are directions, let p and q be vertices such that p is the sole ψ-maximal vertex of H and q is a φ-minimal vertex of H. Then there is a course $p_0, \ldots, p_\nu$ on H from q to p such that (i) for each ι such that $1 \leq \iota \leq \nu$, $\varphi \cdot (p_\iota - p_{\iota-1}) \geq 0$, and (ii) if $\psi \cdot \varphi < 0$, then none of the edges $p_{\iota-1}p_\iota$ lie in direction φ.

Proof: The Lemma clearly holds for two-dimensional polygons. Let H be of dimension λ, and suppose the Lemma holds for polyhedra of all lesser dimensions. Where $E^{\psi\varphi}$ is a two-space containing directions ψ and φ, let π be the orthogonal projection onto $E^{\psi\varphi}$. Then πH is a convex polygon. For any direction ψ', a side of πH which consists of all and only the ψ'-maximal points of πH is the π-projection of the face of H (which may be an edge) consisting of all and only ψ'-maximal points of H. Each vertex of πH is the π-projection of at least one vertex of H.

Let $q^* = \pi q$. On the polygon πH, there is a course $q_0^*, \ldots, q_\mu^*$ satisfying (i) and (ii), where $q_\mu^* = \pi p$, q_0^* is φ-minimal, and either $\pi q = q_0^*$ or $q_0^* \pi q$ is a segment of the φ-minimal side of πH. In the latter case, by inductive hypothesis (i), where q_0 is a vertex of H and $q_0^* = \pi q_0$, there is a course from q_0 to q on the φ-minimal face F_0 of H. The reverse path from q to q_0 satisfies (i) and (ii) because $\varphi \perp F_0$. Now for each $\iota : 0 \leq \iota \leq \mu$, let q_ι be a vertex of H such that $\pi q_\iota = q_\iota^*$. For $1 \leq \iota \leq \mu$, let F_ι be the face of H consisting of all points of H which project onto side $q_{\iota-1}^* q_\iota^*$. Let ψ_ι be the direction in $E^{\varphi\psi}$ in which F_ι (and hence πF_ι) is extreme, and let χ_ι be the direction from $q_{\iota-1}^*$ to q_ι^*. Then $q_{\iota-1}$ is a χ_ι-minimal vertex of F_ι and q_ι is another vertex of F_ι. By inductive hypothesis (i), there is a path $q_{\iota 0}, \ldots, q_{\iota\lambda(\iota)}$ from $q_{\iota-1}$ to q_ι in F_ι, such that where $r_{\iota\kappa} = q_{\iota\kappa} - q_{\iota,\kappa-1}$, $\chi_\iota \cdot r_{\iota\kappa} \geq 0$ for each $\kappa : 1 \leq \kappa \leq \lambda(\iota)$. Since $\psi_\iota \perp F_\iota$, $\psi_\iota \cdot r = 0$. Since ψ_ι and χ_ι are orthogonal directions in $E^{\psi\varphi}$ and φ lies in $E^{\psi\varphi}$, we have that $\varphi = (\chi_\iota \cdot \varphi)\chi_\iota + (\psi_\iota \cdot \varphi)\psi_\iota$. Hence

$$\varphi \cdot r_{\iota\kappa} = [(\chi_\iota \cdot \varphi)\chi_\iota + (\psi_\iota \cdot \varphi)\psi_\iota] \cdot r_{\iota\kappa}$$

$$= (\chi_\iota \cdot \varphi)(\chi_\iota \cdot r_{\iota\kappa}) + (\psi_\iota \cdot \varphi)(\psi_\iota \cdot r_{\iota\kappa}).$$

Since $\chi_\iota \cdot \varphi \geq 0$, $\chi_\iota \cdot r_{\iota\kappa} \geq 0$, and $\psi_\iota \cdot r_{\iota\kappa} = 0$, it follows that $\varphi \cdot r_{\iota\kappa} \geq 0$. Thus combining the courses from q to q_0 and from each $q_{\iota-1}$ to q_ι, we get a course on H from q to p satisfying (i). If $\psi \cdot \varphi < 0$, then no side of πH in the π-projection of this course is oriented in direction φ, and hence no edge in this path on H from q to p is oriented in direction φ. Thus (ii) is satisfied.

3. COROLLARIES

Player k is <u>dictator</u> under g iff for every alternative x there is a strategy s_k^x for k such that for every environment $\underset{\sim}{v}_{-k}$, $g(x, s_k^x \underset{\sim}{v}_{-k}) = 1$. Game form g is <u>dictatorial</u> iff some player is dictator under g.

<u>Condition of Attainability</u>: For every $x \in V$ there is an $s \in \mathcal{S}$ such that $g(x,s) = 1$.

<u>Corollary</u> 1. Let $|V| \geq 3$. If g is straightforward and satisfies Attainability, then on the domain $\mathcal{T}^*$ of the theorem, g is a probability mixture of dictatorial game forms.

<u>Proof</u>: Consider a player k, and let B_x be the set of all utility scales that rank x first. Since B_x has a non-empty interior and the sets $D_k(s_k)$ cover $\underset{\sim}{R}^V$, some $D_k(s_k)$ with interior points must intersect B_x. Let t_k^x be such that $D_k(t_k^x)$ has interior points and intersects B_x, and let $U_k \in D_k(t_k^x)$, for each k. Let $\underset{\sim}{s}^0$ be such that $g(x, \underset{\sim}{s}) = 1$, and let $\underset{\sim}{s}^i = t_i^x \underset{\sim}{s}_{-i}^{i-1}$ for each i, so that $\underset{\sim}{s}^n = \underset{\sim}{t}^x$. Then if $g(x, \underset{\sim}{s}^{i-1}) = 1$, then $g(x, \underset{\sim}{s}^i) = 1$, since x is top in U_i and t_i^x is U_i-dominant for i. Hence $g(x, \underset{\sim}{t}^x) = 1$. We have shown that if g satisfies Attainability, then $g \restriction \mathcal{T}^*$, g restricted to $\mathcal{T}^*$, satisfies attainability.

Now $g \restriction \mathcal{T}^*$ is a probability mixture of unilateral and duple game forms. There are no duple game forms in this mixture, for if, say, a yz-duple scheme were in the mixture and $x \notin \{y,z\}$, then we could never have $g(x, \underset{\sim}{s}) = 1$. Hence

$g|\mathcal{T}^*$ is a probability mixture of unilateral game forms $f_1, \ldots, f_n$. Since for each x, $g(x, \underset{\sim}{t}^x) = 1$, we must have $f_i(x, t_i^x) = 1$; hence each f_i is dictatorial. That proves the Corollary.

Where $U_1, \ldots, U_n$ are utility scales over V, a lottery ρ over V is <u>Pareto optimal ex ante</u> with respect to $U_1, \ldots, U_n$ iff there is no lottery σ over V such that $U_i \cdot \sigma \geq U_i \cdot \rho$ for all i.

<u>Condition of Ex Ante Acceptability</u>: For every $U_1, \ldots, U_n$, there is an $\underset{\sim}{s}$ such that $g(\underset{\sim}{s})$ is Pareto optimal *ex ante* with respect to $U_1, \ldots, U_n$.

Corollary 2. If g is straightforward and satisfies *Ex Ante* Acceptability and $|V| \geq 3$, then g is dictatorial.

The proof is similar to that of Corollary 2 in Gibbard (1976).

NOTES

1. A decision scheme as the term is used in this paper is called an "unrestricted decision scheme" in Gibbard (1976).

2. The random dictatorship is discussed both in Zeckhauser (1973) and in Gibbard (1973).

3. In Gibbard (1976), this is stated informally on p. 10. The theorem as formally stated in that paper deals with schemes which are defined only for ballots which consist of strong orderings. The theorem as stated here is trivially a consequence of the theorem as stated in the earlier paper.

4. Roughly this corollary was suggested by Hugo Sonnenschein for the theorem on strategy-proofness of decision schemes.

5. This property is studied by Zeckhauser (1973); his Theorem V (p. 945) is similar to the second corollary in this paper. Zeckhauser, however, imposes a non-dictatorship condition which is stronger than the one used in this paper, and hence the corollary here is independent of Zeckhauser's theorem. (Cf. Gibbard, 1976, p. 7).

REFERENCES

R. Farquharson, Theory of Voting, New Haven: Yale University Press, 1969.

A. Gibbard, "Manipulation of Voting Schemes: A General Result", Econometrica 41 (1973), 587-602.

A. Gibbard, "Manipulation of Schemes That Mix Voting with Chance", Discussion Paper No. 175, Center for Mathematical Studies in Economics and Management Science, Northwestern University. Revised January, 1976.

J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton: Princeton University Press, 1947.

M. A. Satterthwaite, "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions", Journal of Economic Theory 10 (1975), 187-217.

R. Zeckhauser, "Majority Rule with Lotteries on Alternatives", Quarterly Journal of Economics 83 (1969), 696-703.

R. Zeckhauser, "Voting Systems, Honest Preferences, and Pareto Optimality", American Political Science Review 67 (1973), 934-46.