

## A CHARACTERIZATION OF WEIGHTED VOTING

ALAN TAYLOR AND WILLIAM ZWICKER

(Communicated by Andreas R. Blass)

**ABSTRACT.** A simple game is a structure  $G = (N, W)$  where  $N = \{1, \dots, n\}$  and  $W$  is an arbitrary collection of subsets of  $N$ . Sets in  $W$  are called winning coalitions and sets not in  $W$  are called losing coalitions.  $G$  is said to be a weighted voting system if there is a function  $w: N \rightarrow \mathbb{R}$  and a "quota"  $q \in \mathbb{R}$  so that  $X \in W$  iff  $\sum\{w(x): x \in X\} \geq q$ . Weighted voting systems are the hypergraph analogue of threshold graphs. We show here that a simple game is a weighted voting system iff it never turns out that a series of trades among (fewer than  $2^{2^n}$  not necessarily distinct) winning coalitions can simultaneously render all of them losing. The proof is a self-contained combinatorial argument that makes no appeal to the separating of convex sets in  $\mathbb{R}^n$  or its algebraic analogue known as the Theorem of the Alternative.

For our purposes, a *simple game* (or *voting game*) will be any pair  $(N, W)$  where  $N = \{1, \dots, n\}$  and  $W$  is an arbitrary collection of subsets of  $N$ . (Additional conditions, such as those in [S], will not be imposed.) Much of the intuition for what we do comes from the context in which  $(N, W)$  corresponds to a voting system in which an alternative (such as a bill or an amendment) is pitted against the status quo. Sets in  $W$  are called *winning coalitions* and correspond (intuitively) to those collections of voters sufficient to guarantee passage if they all vote "aye." Sets not in  $W$  are called *losing coalitions*. The study of simple games goes back to the 1944 work of von Neumann and Morgenstern [VM]; some additional basics are laid out by Shapley [S].

The voting game  $(N, W)$  is said to be a *weighted voting game* (or a *weighted system with quota*) if there exists a function  $w: N \rightarrow \mathbb{R}$  and  $q \in \mathbb{R}$  so that

$$X \in W \quad \text{iff} \quad \sum\{w(x): x \in X\} \geq q.$$

The original European Economic Community [B], where France, Germany, and Italy were given four votes each, Belgium and the Netherlands two votes each, and Luxembourg one vote, with a quota for passage of twelve votes out of the seventeen possible votes, is a real world example of a weighted voting system.

A simple game  $(N, W)$  in which  $W$  is a collection of two element sets corresponds, of course, to a graph. If  $(N, W)$  is also weighted then it corresponds

---

Received by the editors August 10, 1990 and, in revised form, February 4, 1991.

1991 *Mathematics Subject Classification*. Primary 90A28; Secondary 90D12, 05C65.

The authors' research was partially supported by NSF Grants DMS-890095 and DMS-9101830.

©1992 American Mathematical Society  
 0002-9939/92 \$1.00 + \$.25 per page

to a threshold graph, and these have been well-studied [W]. In particular, the problem of characterizing exactly when a graph is a threshold graph has been handled by Chvatal and Hammer [CH], Peled [P], and others. The analogous question in our present context, that is, the finding of a simple property that characterizes weighted games among simple games, goes back at least to Isbell in the mid-1950s [IS]. One answer that has been provided (see [HIP]) is the observation that if we identify a coalition with a point in  $n$ -space, then  $(N, W)$  is weighted iff the convex hull of the winning coalitions is disjoint from the convex hull of the losing coalitions. (This identification allows one to appeal to the well-known separation theorem for convex polytopes.) Another answer (again using known results about separating convex sets) is the recent theorem of Einy and Lehrer [EL] for the monotonic case. The drawback to these approaches lies in the extent to which both the properties and proofs make essential use of ideas that seem less than indigenous to either the game-theoretic or voting-theoretic context to which they apply.

Our goal in what follows is to characterize weighted voting in terms of the ways in which coalitions can gain or lose by trading among themselves. The extent to which trading is intrinsic to an analysis of voting has been clear since the introduction by Isbell [IS] in the 1950s of the desirability relation (where  $p$  is defined to be at least as desirable as  $q$  if every winning coalition containing  $q$  but not  $p$  remains winning when  $q$  is traded for  $p$ ). The trading property we introduce here can be motivated in part by the following. One hears in sports of a “trade that helps (or hurts) both teams.” This is certainly possible in team sports like football or basketball, where success depends not just on the individual talents of each player but how these talents mesh together. On the other hand, this is essentially not possible in individual sports (contested in a team setting) like bowling, golf, track, or gymnastics. Sports in the former group correspond to nonweighted systems; those in the latter group to weighted systems. This basic idea is formalized in the following.

**Definition.** A voting game will be called *k-trade robust* if it is never possible to convert a sequence of  $k$  (not necessarily distinct) winning coalitions into a sequence of  $k$  (not necessarily distinct) losing coalitions via a series of trades (of perhaps different size batches of players) between pairs of coalitions.  $G$  will be called *trade robust* if it is *k-trade robust* for all  $k$ .

Our main result is the following.

**Theorem A.** For any voting game  $G = (N, W)$  where  $|N| = n$ , the following are equivalent:

- (1)  $G$  is a weighted system.
- (2)  $G$  is trade robust.
- (3)  $G$  is  $2^{2^n}$ -trade robust.

Notice that a naive checking of (3) is a finite (albeit lengthy) process, whereas a naive checking of either weightedness directly or the intersection of the convex hulls of  $W$  and  $\mathcal{P}(N) - W$  is an infinite process. Moreover, Theorem A actually provides a fairly simple and uniform procedure for showing that certain games are not weighted: one produces a sequence of winning coalitions (often two in the real world, although we show elsewhere [TZ1] that fixed  $k$

will not suffice) and indicates trades among them that convert all of them to losing coalitions. Real world examples, including the U. S. Legislature and the procedure to amend the Canadian Constitution, appear in [TZ2].

We leave the proof that (1) implies (2) in Theorem A to the reader; of course, (2) implies (3) is trivial. Our proof that (3) implies (1) will proceed by inductively constructing the weighting in such a way that at each stage, the unweighted part acts like it is reasonably trade robust and the weighted part behaves as if it were part of a correct weighting. This is formalized in the following.

**Definition.** Suppose that  $(N, W)$  is a simple game where  $N = \{1, \dots, n\}$  and  $W \subset \mathcal{P}(N)$ . If  $A \subset N$  and  $f: A \rightarrow \mathbb{R}$  then we will say that  $f$  is *trade robust for A* iff the following holds:

Whenever  $k \leq 2^{2^{|N-A|}-1}$  and  $(X_1 \cup Y_1, \dots, X_k \cup Y_k)$  and  $(X'_1 \cup Y'_1, \dots, X'_k \cup Y'_k)$  are two sequences of (not necessarily distinct) coalitions satisfying:

- (1) for  $i = 1, \dots, k$ ,  $X_i \cap Y_i = 0$  and  $X'_i \cap Y'_i = 0$ ,
- (2)  $Y_i \subset A$  and  $Y'_i \subset A$ ,
- (3)  $\sum_{i=1}^k (\sum \{f(p) : p \in Y_i\}) \leq \sum_{i=1}^k (\sum \{f(p) : p \in Y'_i\})$ ,
- (4) for every  $p \in N$ ,  $|\{i : p \in X_i\}| = |\{i : p \in X'_i\}|$ ,

then we cannot have  $X_i \cup Y_i$  winning for every  $i$  and  $X'_i \cup Y'_i$  losing for every  $i$ .

It is easy to see that if  $G$  is  $2^{2^n}$ -trade robust then the empty function is trade robust for  $A = 0$ , and if  $f$  is trade robust for  $A = N$  then  $(N, W)$  is weighted. (The former statement requires the observation that if condition 4 above holds, then the  $X_i$  can be converted to the  $X'_i$  by a sequence of trades. This is, of course, completely trivial if the  $X_i$  are disjoint, for one can then associate an ordered pair  $(i, j)$  to each person  $p$  to indicate that they start in  $X_i$  and end up in  $X'_j$ . If the  $X_i$  are not disjoint then condition 4 allows one to add subscripts for the various occurrences of  $p$  and to then treat the  $X_i$  as if they were disjoint.)

With these observations, it is now easy to see that Theorem A is a consequence of the following.

**Theorem B.** Suppose  $G = (N, W)$  is a simple game,  $A \subset N$ , and  $f$  is trade robust for  $A$ . Suppose  $a_0 \in N - A$ . Then there exists  $c_0 \in \mathbb{R}$  so that  $f \cup \{(a_0, c_0)\}$  is trade robust for  $A \cup \{a_0\}$ .

*Proof.* The intuition behind the argument is that when a real number  $c$  fails to be an appropriate choice for the weight of  $a_0$ , it can be classified as failing because it is too “light” or too “heavy.” The proof involves making these distinctions precise, and then proving a sequence of claims, which ultimately establishes that the supremum of the “low failures” is strictly less than the infimum of the “high failures.” This, together with the observation that any failure is either a low failure or a high failure, leaves a nonempty open interval of choices for the desired  $c_0$ .

So, let us agree to call a real number  $c$  a “low failure” if there exist sequences

$$(X_1 \cup Y_1, \dots, X_k \cup Y_k) \quad \text{and} \quad (X'_1 \cup Y'_1, \dots, X'_k \cup Y'_k)$$

showing that  $f \cup \{(a_0, c_0)\}$  is not trade robust for  $A \cup \{a_0\}$  and

$$|\{i : a_0 \in Y_i\}| > |\{i : a_0 \in Y'_i\}|.$$

(Intuitively,  $c$  is a choice of weight that is too low for  $a_0$ , and this is exploited in the witnessing sequences by using  $a_0$  excessively in the winning coalitions.) Similarly, we will say that  $c$  is a “high failure” if there exists such a sequence with the inequality going the other way.

*Claim 1.* Every failure is either a high failure or a low failure.

*Proof.* If  $c$  were neither a high failure nor a low failure, then any sequences of coalitions that showed it to be a failure at all would involve an equal number of occurrences of  $a_0$  among the  $Y_i$  and among the  $Y'_i$ . But then the occurrences of  $a_0$  could all be shifted from the  $Y_i$  and  $Y'_i$  to the corresponding  $X_i$  and  $X'_i$ , with conditions 3 and 4 in the definition above preserved. This would then show that  $f$  is not trade robust for  $A$ , a contradiction.

*Claim 2.* The real number  $c$  is a low failure iff there are sequences witnessing the failure of  $c$  for which  $a_0$  occurs in none of the  $Y'_i$ .

*Proof.* Given any two sequences that witness the fact that  $c$  is a low failure, simply shift each occurrence of  $a_0$  among the  $Y'_i$  to the corresponding  $X'_i$ . At the same time choose an equal number of occurrences of  $a_0$  among the  $Y_i$  and shift these to the corresponding  $X_i$ . Again, it is easy to see that 3 and 4 are preserved, and the resulting sequences are the desired kind of witness to the low failure of  $c$ .

*Claim 3.* The real number  $c$  is a high failure iff there are sequences witnessing the failure of  $c$  for which  $a_0$  occurs in none of the  $Y_i$ .

*Proof.* Analogous to that of Claim 2.

*Claim 4.* If  $c$  is a low failure and  $c' \leq c$ , then  $c'$  is also a low failure.

*Proof.* Any sequences that witness the fact that  $c$  is a low failure will also witness the fact that  $c'$  is a low failure, since condition 3 is preserved by decreasing  $c$  in such a witness.

*Claim 5.* If  $c$  is a high failure and  $c' \geq c$  then  $c'$  is also a high failure.

*Proof.* Analogous to that of Claim 4.

*Claim 6.* No failure can be both a low failure and a high failure.

*Proof.* Assume, for contradiction, that  $c$  is a low failure as witnessed by

$$(*) \quad (X_1 \cup Y_1, \dots, X_k \cup Y_k) \quad \text{and} \quad (X'_1 \cup Y'_1, \dots, X'_k \cup Y'_k),$$

(chosen as in Claim 2) and at the same time a high failure as witnessed by

$$(**) \quad (U_1 \cup V_1, \dots, U_l \cup V_l) \quad \text{and} \quad (U'_1 \cup V'_1, \dots, U'_l \cup V'_l)$$

(chosen as in Claim 3). Thus, we have  $k, l \leq 2^{|N - (A \cup \{a_0\})| - 1}$ . Assume that  $|\{i: a_0 \in Y_i\}| = s$  and  $|\{i: a_0 \in Y'_i\}| = t$ . Consider the sequence of “unprimed” coalitions obtained by repeating each unprimed coalition  $X_i \cup Y_i$  in  $(*)$  a total of  $t$  times and each unprimed coalition  $U_i \cup V_i$  in  $(**)$  a total of  $s$  times, together with the sequence of “primed” coalitions obtained in a similar way from the primed coalitions in  $(*)$  and  $(**)$ . In the combined system,  $a_0$  occurs  $st$  times among the  $Y_i$  and no times among the  $V_i$ , and it occurs  $st$  times among the  $V'_i$  and no times among the  $Y'_i$ . Thus, if we now shift  $a_0$  from the  $Y_i$  to the

corresponding  $X_i$  and from the  $V'_i$  to the corresponding  $U'_i$ , then condition (4) is preserved. It is also easy to see that condition (3) is preserved, since we are reducing each of the two relevant sums by exactly  $s$  times  $t$  times  $c$ . Moreover, the length of the sequences in the combined system is at most

$$2 \left( 2^{2^{|N-(A \cup \{a_0\})|-1}} - 1 \right)^2.$$

This is, of course, just  $2^{2^{|N-A|-1}}$ , and thus we have a contradiction in the form of a witness to the fact that  $f$  is not trade robust for  $A$ .

*Claim 7.* If  $c = 2(2^{2^{|N|}}) \sum \{|f(p)| : p \in A\}$  and  $c' > c$ , then  $c'$  is not a low failure. If  $c' < -c$  then  $c'$  is not a high failure.

*Proof.* Assume for contradiction that  $c' > c$  and  $c'$  is a low failure with witness chosen as in Claim 2. Condition (3) asserts that

$$\sum_{i=1}^k \left( \sum \{|f(p)| : p \in Y_i\} \right) \leq \sum_{i=1}^k \left( \sum \{|f(p)| : p \in Y'_i\} \right);$$

but if  $a_0$  appears among the  $Y_i$  and not among the  $Y'_i$  then

$$\begin{aligned} \sum_{i=1}^k \left( \sum \{|f(p)| : p \in Y_i\} \right) &> c - 2^{2^{|N|}} \sum \{|f(p)| : p \in A\} \\ &\geq 2^{2^{|N|}} \sum \{|f(p)| : p \in A\} \\ &\geq \sum_{i=1}^k \left( \sum \{|f(p)| : p \in Y'_i\} \right). \end{aligned}$$

The other argument is analogous.

*Claim 8.* The low failures are bounded above and the high failures are bounded below.

*Proof.* From Claim 7.

*Claim 9.* If  $c$  is the supremum of the low failures then  $c$  is itself a low failure.

*Proof.* Assume not. Then for some sequence  $c_1 < c_2 < \dots$  converging to  $c$  there exists a single pair of sequences of coalitions that simultaneously shows that each  $c_i$  is a low failure. But then, since we have that for each  $j$ ,

$$\sum_{i=1}^k \left( \sum \{|f(p)| : p \in Y_i\} \right) \leq \sum_{i=1}^k \left( \sum \{|f(p)| : p \in Y'_i\} \right)$$

is true for  $f(a_0) = c_j$ , the inequality still holds when  $f(a_0) = c$ . This shows that  $c$  is a low failure.

*Claim 10.* If  $c$  is the infimum of the high failures, then  $c$  is itself a high failure.

*Proof.* As in Claim 9.

Hence, the low failures make up an interval closed on the right, and the high failures make up an interval closed on the left that is disjoint from it. This leaves a nonempty open interval of values which are not failures.

Added in proof. Our original proof of the characterization theorem, which did use separation of polytopes, is apparently less capable of being mined for additional information. For example, in a forthcoming paper we show how similar ideas can be used to construct an algorithm for determining the range of possible weights for a voter.

#### ACKNOWLEDGMENT

The authors would like to thank Steven Brams, Andreas Blass, and the referee for helpful discussions and suggestions with this and related material.

#### REFERENCES

- [B] S. J. Brams, *Rational politics*, Congressional Quarterly Press, Washington, 1985.
- [CH] V. Chvatal and P. L. Hammer, *Aggregation of inequalities in integer programming*, Ann. Discrete Math. **20** (1977), 145–162.
- [EL] E. Einy and E. Lehrer, *Regular simple games*, Internat. J. Game Theory **18** (1989), 195–207.
- [HIP] P. L. Hammer, T. Ibaraki, and U. N. Peled, *Threshold numbers and threshold completions*, Studies on Graphs and Discrete Programming (P. Hansen, ed.), Ann. Discrete Math., vol. 11, North-Holland, Amsterdam and New York, 1981, pp. 125–145.
- [IS] J. R. Isbell, *A class of majority games*, Quart. J. Math. Oxford Ser. (2) **7** (1956), 183–187.
- [P] U. N. Peled, *Matroidal graphs*, Discrete Math. **20** (1977), 263–286.
- [S] L. S. Shapley, *Simple games: an outline of the descriptive theory*, Behavioral Sci. **7** (1962), 59–66.
- [TZ1] A. Taylor and W. Zwicker, *Threshold hypergraphs, simple voting games, and magic squares*, preprint.
- [TZ2] —, *Weighted voting, multicameral representation, and power*, preprint.
- [VM] J. von Neumann and O. Morgenstern, *Theory of games and economic behavior*, Princeton Univ. Press, Princeton, NJ, 1944.
- [W] D. West, *Parameters of graphs and partial orders: packing, covering and representation*, Graphs and Orders (I. Rival, ed.), Proc. Symposium Banff 1984, Reidel, Dordrecht, 1985, pp. 267–350.

DEPARTMENT OF MATHEMATICS, UNION COLLEGE, SCHENECTADY, NEW YORK 12308