האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

SOCIAL CHOICE AND THRESHOLD PHENOMENA

by

GIL KALAI

Discussion Paper # 279 November 2001

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Social Choice and Threshold Phenomena

Gil Kalai *
Institute of Mathematics
Hebrew University of Jerusalem, Jerusalem, Israel

November 20, 2001

Abstract

Arrow's theorem asserts that under certain conditions every non-dictatorial social choice function leads to nonrational social choice for some profiles. In other words, for the case of non-dictatorial social choice if we observe that the society prefers alternative A over B and alternative B over C we cannot deduce what its choice will be between B and C. Here we ask whether we can deduce anything from observing a sample of the society's choices on the society's choice in other cases?

We prove that the answer is "no" for large societies for neutral and monotonic social choice function such that the society's choice is not typically determined by the choices of a few individuals.

The proof is based on threshold properties of Boolean functions and on analysis of the social choice under some probabilistic assumptions on the profiles. A similar argument shows that under the same conditions for the social choice function but under certain other probabilistic assumptions on the profiles the social choice function will typically lead to rational choice for the society.

1 Introduction

Arrow's impossibility theorem asserts that under certain natural conditions, if there are at least three alternatives then every non-dictatorial social choice gives rise to a non-rational choice function, i.e., there exist profiles such that the social choice is not rational. Arrow's theorem can be seen in the context

*I wish to thank Bezalel Peleg for explaining several issues concerning choice functions and social choice and for several useful discussions. I am also thankful to Peter Fishburn, Yuval Peres, Yossi Rinnot, Ariel Rubinstein and Ran Spiegler for helpful remarks and suggestions. Section 5 was motivated by a question posed by Ran Spiegler.

of Condorcet's "paradox" which demonstrates that the majority rule may result in the society preferring A over B , B over C and C over A. Arrow's theorem shows that such "paradoxes" cannot be avoided with *any* non-dictatorial voting method. It is the general form of Arrow's theorem, which can be applied to general schemes for aggregating individual rational choices, that made it so important in economic theory.

For certain voting methods, the social choice may become chaotic as the number of alternatives increases. McGarvey (1953) appears to have been the first to show that for every asymmetric relation R on a finite set of candidates there is a strict-preferences (linear orders, no ties) voter profile that has the relation R as its strict simple majority relation. This implies that we cannot deduce the society's choice between two candidates even if we know the society's choice between every other pair of candidates. Saari (1989) proved that the plurality method gives rise to *every* choice function for sufficiently large societies. This implies that knowing the outcome of the plurality choice for several examples, where each example consists of a set S of alternatives and the chosen element $c(S)$ for S, cannot teach us anything about the outcome for a set of alternatives which is not among the examples we have already seen.

In this paper we strengthen the condition that the social choice function is not dictatorial and demand (in a technical sense that will be explained later) that the power of individuals be small. As a result we will obtain a much stronger conclusion regarding the resulting choice functions.

We will always refer to the number of alternatives as m and the number of individuals in the society as n. In the first case considered in Section 3 the social choice function is an asymmetric relation $R = F(R_1, R_2, \ldots, R_n)$ on pairs of alternatives which depend on the preference relations R_i of the individuals in the society. We will also assume that the social choice function is neutral and monotone. We obtain the result that every asymmetric relation is in the range of the social choice function. This is precisely the result of McGarvey's theorem for the majority rule.

Given a set X of m alternatives, a choice function c is a mapping which assigns to nonempty subsets S of X an element $c(S)$ of S. A *rational* choice function is one for which there is a linear ordering on the alternatives such that $c(S)$ is the maximal element of S according to that ordering.

In the second more general case considered in Section 4 the social choice function is a choice function $c = F(R_1, R_2, \ldots, R_n)$ on the set of alternatives. Let $c_k(S)$ be the (rational) choice of the k-th individual for the set S. In other words, $c_k(S)$ is the maximal element of S according to the order relation R_k. We assume a strong for of the Independence of Irrelevant Alternatives that we call "Independent of Rejected Alternatives (IRA)". This condition asserts that $c(S)$ depends on the individual choices for the set S, i.e., that $c(S)$ is a function of $c_1(S)$, $c_2(S)$, $\ldots, c_n(S)$. We also require that the social choice function is neutral and monotone. In this case the result we are seeking is that all possible choice functions are within the range of the social choice function. Precisely the result of Saari for the plurality rule.

Some voting methods, such as the Borda rule, the aggregated social choice given a set S of alternatives depends on the actual preferences of the individuals on S and not only on their choices for S. In the Borda rule each individual ranks the elements of S by the numbers $1, 2, \ldots, |S|$ and $c(S)$ is the element for which the sum of the individual rankings is minimal. When we allow such a dependence there are social choice functions based on the Condorcet winners among pairs which give rise to classes of choice functions whose size is only exponential in a quadratic function in m, much smaller than the number of all choice functions which is double exponential in m. It follows from our results that no smaller class of choice functions is possible. For the Borda rule the size of the class of social choice that arise is $O(N^3)$.

The notion of the "power" of an individual within a social choice function is based on the *Shapley value* from game theory. We require that the power of individuals be small which is automatically the case when the number of individuals is large and all individuals have the same power. We define a social choice function as *weakly anonymous* if it is invariant under a transitive group of permutations of the individuals.[1] For example, an electoral voting system such as that in the U.S. in which all states have the same number of voters and the same number of electoral votes is weakly anonymous but not anonymous. For weakly anonymous social choice functions the power of every individual is identical. Another case where the power of every individual is small is when the individuals are classified by "types", two individuals of the same type are indistinguishable and there are many individuals of each type.

One consequence of our results is related to the outcome of social choice functions when individual preferences are restricted. Maskin (1995) and Dasgupta and Maskin (1997) considered the case in which the individual orderings are restricted to a set T of orderings on the alternatives. They proved that if the majority rule (between pairs of alternatives) leads to non-rational preferences for the society under this restriction, then non-rational outcomes cannot be avoided when individual preferences are restricted to T

[1] A group Γ of permutations of $\{1, 2, \ldots, n\}$ is transitive if for every i and j there is a permutation $\pi \in \Gamma$ such that $\pi(i) = j$.

for a large class of social choice functions. It follows from our result that non-rational outcomes cannot be avoided for neutral monotone social choice functions when the power of individuals is sufficiently small.

The proofs are based on threshold properties of Boolean functions and on an analysis of the social choice under some probabilistic assumptions on the profiles. There is a large literature on the probability of non-rational outcomes in voting schemes when individual preferences are uniform and independent and under certain other probabilistic assumptions as well, (see, for example, Gehrlein (1997)).

The relevant facts about threshold phenomena for monotone Boolean functions are described in Section 2 and in more details in Appendix A. Their application in demonstrating the chaotic behavior of social choice is presented in Sections 3 and 4. In Section 5 we use the same results on the threshold behavior of Boolean functions to point out some probabilistic situations in which a rational outcome is expected. This is the case when individual choices are biased towards some fixed order relation and certain other cases when individual choices are positively correlated.

2 Choosing between two alternatives - the sharp threshold phenomenon

2.1 The threshold interval

In this section I present the main results concerning monotone Boolean functions that will be used in this paper. These will serve as technical tools for our purposes although they appears also to be of interest within theoretical economics and game theory. We will give a short and somewhat informal description here leaving the more technical details to Appendix A.

Consider a social choice function between two alternatives a and b for a society with n individuals. We will represent the individual choices by Boolean variables $x_1, x_2, \ldots, x_n$, where $x_k = 0$ if the k-th individual prefers alternative a on alternative b and $x_k = 1$ if he prefers b on a. We will represent the social choice by a Boolean function $f(x_1, x_2, \ldots, x_n)$, where $f(x_1, x_2, \ldots x_n) = 0$ if the society prefers alternative a on alternative b and $f(x_1, x_2, \ldots x_n) = 1$ if the society prefers b on a. We will define f to be monotone if whenever $y_i \geq x_i$, $i = 1, 2, \ldots, n$, $f(y_1, y_2, \ldots, y_n) \geq f(x_1, x_2, \ldots, x_n)$.

Suppose now that each variable x_k is chosen at random to be '1' with probability p and '0' with probability $1 - p$. In other words, the k-th individual prefers b on a with probability p. Suppose also that these choices are independent. Denote by $\mathbf{P}_p(f)$ the probability that f will equal '1', i.e., that the society will prefer b on a.

If the function f is monotone and not constant, then the value of $\mathbf{P}_p(f)$ is a strictly monotonic continuous function of p in the interval $[0,1]$. The more likely it is that individuals prefer b to a the more likely that the society does as well.

Let ϵ, $1/2 > \epsilon > 0$, be a real number. Since $\mathbf{P}_p(f)$ is a strictly monotone and continuous function of p there is a unique value of p denoted by p_1 such that $\mathbf{P}_{p_1}(f) = \epsilon$. There is also a unique value of p denoted by p_2 such that $\mathbf{P}_{p_2}(f) = 1 - \epsilon$.

The interval $[p_1, p_2]$ is called a *threshold interval* and its length $p_2 - p_1$ is denoted by $I_\epsilon(f)$. The value p_c at which $\mathbf{P}_{p_c}(f) = 1/2$, is called the *critical probability* for f.

2.2 The Shapley value

There are several measures of the power of the k-th individual within the social choice function given by f. We will use the Shapley value $s_k(f)$ as a measure of power. Here f is considered to me a cooperative n-player simple game. In Appendix A we will reproduce a definition of the Shapley value which suits our purposes. Recall that $s_1(f) + s_2(f) + \cdots + s_n(f) = 1$. If f is invariant under some transitive group Γ of permutations on $[n] = \{1, 2, \ldots, n\}$, then $s_k(f) = 1/n$ for every k.

2.3 The main result concerning thresholds

The main result used here asserts that if the power of every individual is small then the threshold interval must also be small!

Theorem 2.1. *Consider a monotone Boolean function f on n variables. For every $\epsilon, \delta > 0$ there exists $\gamma > 0$ such that if the Shapley value $s_i(f) \le \gamma$ for every $i = 1, 2, \ldots n$, then $|I_\epsilon(f)| \le \delta$.*

3 Social preferences and thresholds

3.1 The setting

Consider a social choice function which, given a profile of n order relations R_i, $i = 1, 2, \ldots, n$ on m alternatives, yields an asymmetric relation R for the society. Thus $R = F(R_1, R_2, \ldots R_n)$ and F is the social choice function. aR_ib indicates that the i-th individual prefers alternative a over alternative

b. aRb indicates that the society prefers alternative a on alternative b. The social preferences are not assumed to be transitive.

The principal condition (I) of Independent of Irrelevant Alternatives, states that for every two alternatives a and b the set $\{i : aR_ib\}$ determines whether aRb. The social preference between a and b can thus be described by a Boolean function $f_{(a,b)}$ of n variables $x_1, x_2 \ldots x_n$ as follows: Set $x_i = 1$ if aR_ib and $x_i = 0$ otherwise. In addition, let aRb if and only if $f_{(a,b)}(x_1, \ldots x_n) = 1$. The other standard (Pareto) assumption (P) states that $f_{(a,b)}(0, 0, \ldots, 0) = 0$ and $f_{(a,b)}(1, 1, \ldots, 1) = 1$. We will consider the case in which the number of alternatives m is fixed although the size of the society can be arbitrarily large.

3.2 Further assumptions

Unrestricted domains

(U) The social choice function is defined for an arbitrary rational profile of individuals.

Neutrality and weaker conditions:

Neutrality refers to the invariance of the social choice under permutations of the *alternatives*. Consider the following conditions:

(N1) (Neutrality) The social choice is invariant under all permutations of the alternatives.

(N2) If you replace the preference relation between a and b for all individuals, then the social preference between a and b is reversed as well. In other words, $f_{(a,b)}$ satisfies $f_{(a,b)}(1 - x_1, 1 - x_2, \ldots 1 - x_n) = 1 - f_{(a,b)}(x_1, \ldots x_n)$.

(N3) (Weak neutrality) The social choice is invariant to a transitive group of permutations of the alternatives.

(N4) (Balance) For every two alternatives a and b, $\mathbf{P}(a > b) = 1/2$. Here $P(a > b)$ is the probability that $f_{(a,b)}(x_1, \ldots x_n) = 1$ if $x_1, \ldots x_n$ are chosen at random according to the uniform distribution on the 2^n 0-1 sequences of length n.

(N5) (Weak balance) For every two alternatives a and b, $1/10 \leq \mathbf{P}(a > b) \leq 9/10$.

Clearly (N1) implies N(2) and N(3) and (N2) and (N3) both imply (N4) which implies (N5). (The numerical values 1/10 and 9/10 can be replaced by arbitrary ρ and $1-\rho$ for a fixed small real number ρ.)

Anonymity and weaker conditions

Anonymity refers to the invariance to permutations of the individuals. Consider the following conditions:

(A1) (Anonymity) The social choice is invariant to permutations of the individuals

(A2) (Weak Anonymity) The social choice is invariant to a transitive group of permutations of the individuals.

(A2.5[r)] (Replicated individuals) The individuals are divided into types, and there are at least r from each type. The social choice is symmetric on each type.

(A3)[ϵ] (Diminishing Power of Individuals) For every a and b, the *Shapley value* of every individual for $f_{(a,b)}$ is smaller than ϵ.

Clearly, (A1) implies (A2) and (A2.5) for every ϵ, (A2) implies (A3)[ϵ] when $n > [1/\epsilon]$, and (A2.5)[r] implies (A3)[ϵ] when $r > [1/\epsilon]$.

Monotonicity

(M) The function $f_{(a,b)}$ is monotone. This means that if $f_{(a,b)}(x_1, \dots x_n) = 1$ and if $y_i \geq x_i$ for every i then $f_{(a,b)}(y_1, \dots y_n) = 1$ as well.

3.3 The result

Theorem 3.1. *Let X be a set of m alternatives and let R be an arbitrary asymmetric relation on X. There is a probability distribution $\nu = \nu_R$ on the space of orderings of the alternatives with the following property: For every real number $\delta > 0$ there exists $\epsilon = \epsilon(m, \delta) > 0$ such that for every social choice function which satisfies conditions (I), (P), (U) (N4), (A3)[ϵ] and (M), if every individual makes the choice at random and independently according to ν, then for every pair of elements a,b the social choice between a and b coincides with R with probability of at least $1 - \delta$.*

Proof: Consider weights w_π of the $m!$ ranking of the alternatives which sum to one and for two alternatives a, b let $p(a,b) = \sum\{w_\pi : a >_\pi b\}$.

According to the theorem of McGarvey (1953) weights exist such that if aRb then $p(a,b) > 0.5$. In fact, according to Stearns (1959) we can realize R by a majority of m voters and therefore $p(a,b) \geq 1/2 + 1/m$ if aRb and $p(a,b) \leq 1/2 - 1/m$ if bRa. Next, for every member of the society consider a random ordering of the alternatives in which they are ordered according to π with probability w_π.

By the threshold theorem the probability that the social choice will agree with R on a pair (a,b) tends to one as the number of individuals tends to infinity. In order for aRb to hold with probability of at least $1-\delta$ we require that $1/m \leq K \log(1/\delta)/\log(1/\epsilon)$, i.e. that $\log(1/\epsilon) \leq Km \log(1/\delta)$.

If the social choice satisfies the weak anonymity property, namely it is invariant under a transitive group of permutations of the individuals, then we require that $n > exp(Km \log(1/\delta))$.

If we desire that with probability of at least $1-\delta$ the social choice will agree with R on all pairs we require that $\log(1/\epsilon) \leq Km \log(\binom{m}{2}/\delta)$. □

Remark: Instead of using random profiles we can simply replicate sufficiently many times each individual as prescribed by McGarvey's theorem.

Corollary 3.2. *Under the conditions of Theorem 3.1 the class of preference relations for the society is the class of all asymmetric relations.*

Remark: Note that in order to realize all asymmetric relations we need that the number of individuals be exponential in the number of alternatives. It seems possible that a polynomial (or even close to linear) number of individuals in terms of the number of alternatives would suffice.

The following corollary can also be derived from the proof of Theorem 3.1:

Corollary 3.3. *For every natural numbers n and m and a real number $\delta > 0$ there exists $\epsilon = \epsilon(n,\delta) > 0$ such that the following holds: Let T be a subset of linear orders on m alternatives such that for the majority rule on n individuals we can restrict the profile to T and obtain a preference relation R for the society. Then for every social choice which satisfies conditions (I), (P), (N4), (A3[ϵ]) and (M) there is a probability distribution ν on T such that if every individual makes the choice randomly according to ν then for every pair a and b, the social preference for the pair (a,b) coincide with R with probability $> 1-\delta$.*

The following Corollary is in the spirit of theorems by Maskin (1995) and Dasgupta and Maskin (1997):

Corollary 3.4. *There exists $\epsilon = \epsilon(m) > 0$ with the following property: Let T be a set of linear orders on m alternatives. If the majority rule can yield a non-rational outcome for profiles restricted to T, then every social choice which satisfies conditions (I),(P),(N4), (A3[ϵ]) and (M) leads to a non-rational outcome for some profile restricted to T.*

Dasgupta and Maskin (1997) proved this statement without assuming monotonicity under anonymity and neutrality when there are sufficiently many individuals in the society.

4 The chaotic behavior of social choice

4.1 The setting

Given a set X of m alternatives recall that a choice function c is a mapping which assigns to nonempty subsets S of X an element $c(S)$ of S. In the case of a a rational choice function there is a linear ordering on the alternatives such that $c(S)$ is the maximal element of S according to that ordering. Let $P_+(X)$ denote the family of non-empty subsets of X.

The social choice functions considered in this section are of the form $c = F(R_1, R_2, \ldots, R_n)$ where c is a choice function on X which depends on the profile of individual preferences $R_1, \ldots R_n$. Independence of Irrelevant Alternatives asserts that $c(A)$ may depend only on the preference relations restricted to the set A. We require the stronger property Independence of Rejected Alternatives:

(IRA) (Independence of Rejected Alternative (IRA)) $c(S)$ is a function of $(c_1(S), c_2(S), \ldots c_n(S))$

Therefore we can write $c(S) = F_S(c_1(S), c_2(S), \ldots c_m(S))$.

(P) (Pareto) If $s(S) \in \{c_1(S), c_2(S), \ldots, c_n(S)\}$.

4.2 Further assumptions

We will make the following additional assumptions:

(U) (Unrestricted domains) The social choice function is defined for arbitrary rational profiles of the individuals.

(N1') (Neutrality) The social choice is invariant under permutations of the alternatives.

(A2') (Weak anonymity) The social choice is invariant under a transitive group of permutations of the individuals.

(M') (Monotonicity) The function $c(S)$ is monotone in the following sense: If $c(S) = s$ and $c_i(S) = t$ for $t \neq s$ then changing the choice of the i-th individual from t to s will not change $c(S)$.

4.3 The result

Theorem 4.1. *Let c_0 be a choice function on $P_+(X)$ where X is a set of m alternatives. There is a probability distribution ν on the space of orderings of the alternatives such that the following holds: For every positive real number $\delta < 1$, there exists $N = N(\delta, m)$ such that for every social choice function which satisfies conditions (IRA) (P) (U') (N1'), (A2') and (M'): If the number of individuals n is larger than N and if every individual makes the choice randomly and independently according to ν, then for the social choice $c(S) = c_0(S)$ with probability of at least δ.*

The proof is indicated in Appendix B.

Corollary 4.2. *There exists $N = N(m)$ such that when the number n of individuals is larger than $N(m)$, every choice function on $P_+(X)$, where X is a set of m alternatives, is in the range of every social choice which satisfies conditions (IRA), (P), (U') (N1'), (A2') and (M').*

Again, in the spirit of Dasgupta and Maskin's theorem we have the following:

Corollary 4.3. *There exists $N = N(m) > 0$ with the following property: Let U be a set of linear orders on m alternatives. If for the plurality rule there is a profile restricted to U which leads to a choice function c for the society, then when $n > N(m)$ this is the case for every social choice function which satisfies conditions (IRA), (P), (N1'), (A2') and (M').*

5 Probabilistic assumptions which lead to rational outcomes

We will now consider an additional application of the threshold phenomenon which in this case leads to the conclusion that the social choice *is* rational. There are several papers which show that under certain probabilistic conditions voting paradoxes are rare (see, for example, Tangian(2000). Here we would like to produce results which apply to arbitrary social choice functions.

We will consider two simple models. In the first, the profiles are biased towards some fixed order. Let $q < 1$ be a fixed real number.

For an ordering π of $\{1, 2, \ldots, m\}$ let $i(\pi)$ be the number of inversions in π. Here, $\{i, j\}$ is an inversion if $i < j$ but $\pi(i) > \pi(j)$.

- The choices of each individual are rational. The probability of an ordering π is proportional to $q^{i(\pi)}$.

It is easy to see that for the majority rule, if the size of the society tends to infinity, then the outcome for both these rules will tend to the ordering $1 < 2 < 3 < \cdots < m$.

Theorem 5.1. *For every social choice function that satisfies (P),(I),(N4), (A2) and (M), as the number of individuals tends to infinity the probability that iRj when $i < j$ tends to 1.*

Proof: In the rules described above the probability that an individual will prefer i to j when $i < j$ is $(1 + q)/2 < 1/2$. By the threshold theorem under our conditions the probability that the society will prefer i to j tends to zero as the number of individuals tends to infinity. (Of course, for all the results of this section condition (A2) and the assumption that the number of individuals tends to infinity can be replaced by (A3)[ϵ] and the assumption that ϵ tends to zero.)

Remark: The result and proof extends to the case in which the individual preferences are arbitrary asymmetric relations (not rational) which are biased towards a fixed order relation.

The second model describes a situation in which the individual choices are not biased but are correlated. The model is a simple variation of the Ising and Potts models from statistical physics. Let $G =< V(G), E(G) >$ be a graph.

- The individuals are placed on the vertices v of G. The preference relation of the individual located at v is rational and is described by the ordering π_v. The probability for a set of orderings $\{\pi_v : v \in G\}$ is proportional to:
$$\prod_{\{v,u\} \in E(G)} \exp(t \cdot i(\pi_v \pi_u^{-1})).$$

Theorem 5.2. *(1) Let G be the complete graph with n vertices, or a graph of a rectangular grid in the plane. There exists T_0 such that whenever $t > T_0$, for every social choice function that satisfies (I), (P), (N4), (A3) and (M), the probability that the outcome will be rational tends to 1 as the number of individuals tends to infinity.*

Proof: It follows from standard facts about the Potts model on graphs that for the cases described above that the distribution for the order relations on the vertices of the graph will be biased towards a single order relation with high probability. This property allows us to use the threshold property for social choice as before. □.

Remarks: 1. It would be interesting to relate the phenomenon described in this section to economic models which attempts to predict the outcome of the aggregated actions of individuals and, in particular, to models concerning equilibrium.

2. The Ising and Potts models are related to dynamical processes in which each voter's choice is influenced by the choices of its neighbors.

3. In some situations when we allow individual preferences to be irrational increasing the correlation between individual choices result in them being less rational and the social choice more rational.

6 Taking rejected alternatives into account

This section describes a crucial example suggested by Bezalel Peleg. First, recall how we can progress from choices on pairs to a choice correspondence: Given an asymmetric binary relation R on the set of alternatives let $c(S)$ be the set of elements y of S such that the number of $z \in S$ such that yRz is maximal. In other words, when we consider the directed graph described by the relation we choose the vertex of maximal out-degree.

Let $\mathcal{R}$ denote the class of rational choice functions and $\mathcal{B}$ denote the class of choice correspondences obtained from binary relations R as just described. Consider also the class $\mathcal{B}'$ of choice functions obtained from $\mathcal{B}$ by choosing a single element in $c(S)$ according to some fixed order relation on the alternatives. The number of choice functions in $\mathcal{B}'$ and the number of choice correspondences in $\mathcal{B}$ is exponential in $\binom{n}{2}$.

Now describe a social choice function as follows: aRb if a majority of the society prefers a to b.

In this case the social choice of S does not depend solely on the individual choices for S but also on the preferences among pairs of elements in S.

When the individual choices are rational then the social choice still belongs to the class $\mathcal{B}$ (or $\mathcal{B}'$). In this case the choice from S is simply those elements of S which are Condorcet winners against the maximal number of other elements in S. In this example the social choice for a set S is typically large but this apparently be corrected by various methods of "tie breaking".

In these examples the size of the resulting classes of choice functions is exponential in a quadratic function of m^2. It is much smaller than the number of all choice functions which is double exponential in m.

The Borda rule can be analyzed by a similar consideration, see Kalai (2001). For this rule $c(A)$ is determined as follows: For each alternative $a \in A$ let $r_i(a)$ be the number of individuals who ranked a in the ith place (among the elements of A). Let $r(a) = \sum i \cdot r_i(a)$. The chosen element by the society $c(A)$ is the element of a with the minimal weight.

Another way to describe the Borda rule is as follows: First construct a directed graph (with multiple edges) with A as the set of vertices by introducing an edge from a to b for every individual that prefers a to b. Next, define (as before) $c(A)$ as the vertex with maximal outdegree.

It is easy to prove that the number of choice functions that arise in this way is at most exponential in m^3. (The choice function can be recovered from the sign patters of (less than) $2^m \cdot m^2$ linear expressions in m^2 real variables. (See, Kalai (2001a).)

To summarize, the size of classes of choice functions that arise from a social choice function such that $c(S)$ may depend on the individual preferences of the elements of S is at least exponential in m^2 and this bound is sharp.

7 Extending the scope of the main theorems

7.1 Social choice with no small set of decisive individuals

Condition (A3), which requires that the Shapley value of every individual is small can be replaced by a weaker condition (A4) which allows for the existence of some powerful individuals as long as there are no small sets of individuals with decisive power.

(A4)$[\delta, r]$ (Non-existence of a small set which typically decides outcomes) Here $\delta < 1$ and r is a natural number. For every a and b the sum of the Shapley values for every r players in $c(a, b)$ is at most $1 - \gamma$.

Clearly, (A3)$[\epsilon]$ implies (A4) $[1 - \epsilon, [1/\epsilon] - 1]$.

Theorem 3.1 remains true if we replace (A3)$[\epsilon]$ in which the size ϵ depends on the number of alternatives m and on the probability γ, by (A4)$[\delta, r]$, in which $\delta < 1$ is fixed and how large r is depends on m, γ and δ.

7.2 Weakening and removing the monotonicity assumption

Monotonicity is a natural condition to demand but is often unrealistic. The question of whether this condition can be removed is therefore of interest. Intuitively, it appears that non-monotonic methods of aggregating individual choices will yield chaotic outcomes.

Let p be fixed and let $I_p^+(A)$ be the probability that toggling the value of x_k from 0 to 1 will change the value of f from 0 to 1 and let $I_p^-(A)$ be the probability that toggling the value of x_k from 0 to 1 will change the value of f from 1 to 0. Let $I_p^+(A) = \sum_{k=1}^n I_p^+(A)$ and $I_p^+(A) = \sum_{k=1}^n I_p^+(A)$.

(M2)[θ] **Weak monotonicity:** For some constant $\theta > 1$ and for every p

$$I_p^+(A) \geq \theta I_p^-(A).$$

The assertion of Theorems continues to hold if (M) is replaced by (M2)[θ] for a fixed $\theta > 1$.

When we remove the monotonicity condition altogether, then the assertions of Theorems 3.1 and 4.1 are no longer valid. To see this consider a case in which n is odd and $f_{(a,b)}$ is a if an odd number of individuals prefer a and $f_{(a,b)} = b$ if an odd number of individuals prefer b.

However, the assertions of Corollaries 3.2, 3.3 3.4, 4.2 and 4.3 probably remain valid. To this end we would like to pose the following conjecture: [2]

Conjecture 7.1. *If the social choice satisfies conditions (A2) (weak anonymity) and (N3) (neutrality) then for every asymmetric relation R the probabilities p_R that for random uniform profiles the social choice will yield R is bounded away from zero as n tends to infinity.*

8 Comments

8.1 Learnability

The motivation for the results described in this paper were derived from my attempts to study the learnability properties of classes of choice functions that describe individual and collective choice in theoretical economics.

In Kalai (2001a) it is argued that the choice functions of individuals and even several interacting individuals which appear in theoretical economics

[2] This conjecture has now been proved by Friedgut and the author

are "learnable" from a "small" number of examples. The notion of PAC-learnability is used there to analyze learnability. Rational choice functions are statistically learnable from a number of examples which is proportional to the number of alternatives. Various classes of choice functions are shown to be learnable from a number of examples which is bounded by a polynomial in the number of alternatives. A closely related (much simpler) feature of rational choice functions is that they can be *described* by a small number of examples. A specific order relation is determined completely by $n - 1$ examples.

We know that a non-dictatorial social choice cannot (under Arrow's assumptions) guarantee rational choice. Can a social choice mechanism guarantee that the class of choice functions that arise for the society will be statistically learnable from a number of examples which is polynomial in the number of alternatives? If we cannot guarantee a rational social choice can we at least guarantee that the class of choice functions that represent the society's choices will allow parsimonious description?

Our results support the conjecture made in Kalai (2001a) that no such social choice mechanism exists if the social choice genuinely depends on a large number of individuals and the society's choices depend only on the individual choices.

8.2 On threshold phenomena

The main mathematical tool used in the paper is the threshold behavior of social choice functions for two alternatives. The threshold results enable us to demonstrate that "Bad" behavior and "good" behavior for the majority (or plurality) rule is carried over to arbitrary monotone social choice functions when the power of individuals diminishes.

It would be of interest to study threshold phenomena in economic systems for their own sake and not just as mathematical tools. The possibility that certain economic systems demonstrate sharp transitions between two (or more) types of behavior is intuitively appealing, and the results concerning threshold intervals may be useful in analyzing such situations.

9 Discussion

We have considered social choice functions in which the power of individuals diminishes and more general "genuine" social choice functions in which no small set of individuals has decisive power. Our notion of "power" is based on the Shapley value. Examples where the power of the individuals diminishes is when there is a large number of them and the social choice function is invariant under a transitive group of permutations of the individuals.

In the first case we considered the social choice function yield an asymmetric relation on the alternatives and our main result asserts that.

[1] Every genuine social choice function which is neutral and monotone leads to the class of all asymmetric relations for large societies.

Next we considered more general social choice functions that yield a choice function on the alternatives. We demonstrated:

[2] Every social choice function which is weakly anonymous, neutral and monotone in which the society's choice is a function of the individual choices leads to the class of all choice functions for large societies.

It follows that the class of asymmetric relations and the choice functions that are obtained in these two cases cannot be learned or described based on a few examples. Based on the society's choices for several sets of alternatives, we are unable to deduce what the choice will be for any other set of alternatives.

Another implication of our results is the following:

[3] There is a sharp difference between classes of choice functions that can arise from voting schemes when the society's choice depends on the individual choices and those when the society's choice depends on the individual preferences for the available alternatives.

Recall that the Independence of Irrelevant Alternatives is the crucial assumption underlying Arrow's theorem. If the social choice takes into account the individual preferences over all possible alternatives, not only the available ones, then non-dictatorial social choice functions are possible. The assumption of Independence of Rejected Alternatives (IRA) is of a similar nature: Can the society's choice for an available set S of alternatives take into account the preference relations of individuals for the elements of S? or only the element chosen by each individual? In the latter case, under our assumptions the class of choice functions that represent the society's choices is the class of all choice functions hence it has double exponential size in terms of the number of alternatives. However, when the social choice can depend on individual preferences the resulting class of choice functions may be rather small and at the minimum will be exponential in a quadratic function of the number of alternatives.

One implication of our proof is closely related to the results of Maskin and Dasgupta:

[4] If a certain restriction T on the individual preferences is sufficient to guarantee rational outcomes for a certain neutral social choice function when the power of every individual is sufficiently small (namely smaller than some function of the number of alternatives), then restricting the individual preferences to T will lead to rational outcomes for the majority rule.

Our proofs were based on an analysis of the social choice function under certain probabilistic assumptions and the threshold phenomenon for Boolean functions. Using similar reasoning we derived the following:

[5] For certain probabilistic models in which individual choices are either biased towards a fixed ordering or positively correlated, the following holds: For social choice functions in which the power of individuals is diminishing the social choice will almost certainly be rational.

In summary, the "bad news" is that under fairly mild conditions every social choice mechanism leads to all possible choice functions when the society is large. In particular, the resulting class of choice functions is highly chaotic. The "good news" is that under certain probabilistic assumptions the outcome of every social choice function in which the individual's power diminishes will be rational with a high probability.

The results of this paper therefore suggest, perhaps unsurprisingly, that the type of interaction between the choices of the individuals and the probabilistic consequences of these interactions may be more important to the aggregate outcome than the combinatorial mechanism used for the aggregation.

For large societies and diminishing power of the individuals not only is there no combinatorial mechanism to guarantee rational outcomes, as implied by Arrow's theorem, but even learnable (or predictable) outcomes cannot be guaranteed. Indeed, under some probabilistic assumptions on individual choices we can expect chaotic behavior for the society's choices. Under certain other probabilistic assumptions the social choice will "invisibly" lead to rational outcomes.

Understanding the conditions under which these two types of behavior arise, the transition between them and possible intermediate situations remains an interesting problem for further research.

References

[1] Ben-Or, M. and N. Linial (1990), Collective coin flipping, in *Randomness and Computation* (S. Micali, ed.), Academic Press, New York, pp. 91–115.

[2] Bollobas, B., G. Grimmett and S. Janson (1996), The random-cluster model on the complete graph. Probab. Theory Related Fields 104, 283–317.

[3] Dasgupta, P.S., and E. Maskin, (1997), "The Robustness of Majority Rule", mimeo, Department of Economics, University of Cambridge.

[4] Erdös, P. and L. Moser (1964), On the representation of directed graphs as unions of orderings, Publ. Math. Inst. Hung. Acad. Sci., 9 (1964), 125-132

[5] Fishburn, P(1973), The Theory of Social Choice, Princeton University Press.

[6] Friedgut, E. and G. Kalai (1996), Every monotone graph property has a sharp threshold, *Proc. Amer. Math. Soc.* **124**, 2993–3002.

[7] Grimmett, G. (1989), *Percolation*, Springer-Verlag, Berlin.

[8] Kahn, J., G. Kalai and N. Linial (1988), The influence of variables on boolean functions, *Proc. 29-th Ann. Symp. on Foundations of Comp. Sci.*, 68–80.

[9] Kalai, G. (2001a) Learnability and rationality of choice, preprint.

[10] Kalai, G. (2001b) A Fourier-Theoretic Perspective for the Condorcet Paradox and Arrow's theorem, to appear, Adv. Appl. Math.

McCoy B. M., and T. T. Wu (1973), The two-dimensional Ising model, Harvard Univ. Press, Cambridge, Mass.

[11] Maskin, E. (1995), Majority rule, social welfare functions and game forms, In K. Basu, P.K. Pattanaik and K. Shuzumura, eds. *Choice, Welfare and development*, Clarendon Press, Oxford.

[12] McGarvey D.C. (1953), A theorem on the construction of voting paradoxes, Econometrica 21, 608-610.

[13] Owen, G. (1988), Multilinear extensions of games. The Shapley value, 139–151, Cambridge Univ. Press, Cambridge, 1988.

[14] Russo, L. (1982), An approximate zero-one law, Z. Wahrsch. werw. Gebiete, 61 (1982), 129-139.

[15] Saari, Donald, G. (1989), A dictionary of voting paradoxes, J. Econ. Th. 48. 443-475.

[16] Saari, Donald, G.(2000), Mathematical structures of voting paradoxes II, positional voting, Economic Theory 15, 55-102.

[17] Sen, Amartya (1986) Social choice theory, Ch. 22, Vol III of Arrow and Intriligator(1986), 1073-1182.

[18] Stearns, R. (1959), The voting problem, American Mathematical Monthly, 66, 761-763.

[19] Talagrand, M. (1994), On Russo's approximate zero-one law, *Ann. of Prob.* **22**, 1576–1587.

[20] Tangian, A. S. (2000), Unlikelihood of Condorcet's paradox in a large society, Social Choice and Welfare 17, 337-365.

10 Appendix A: Thresholds, influences and the Shapley value

Influences, Russo's Lemma and the Shapley value

Consider a monotone non-constant Boolean function $f(x_1, x_2, \dots, x_n)$.

The *influence* of the variable k on the Boolean function f, denoted by $I_k^p(f)$, is the probability that toggling the value of x_k will change the value of f. The total influence $I^p(f)$ equals $\sum I_k^p(f)$.

Russo's lemma (see Grimmett (1989) asserts that

$$\frac{d\mathbf{P}_p(f)}{dp} = I^p(f).$$

By Russo's lemma, if the total influence in the threshold interval is large then this threshold interval is small.

Let $\mathbf{P}_{(p_1,p_2,\dots,p_n)}(f)$ be the probability that $f(x_1, x_2, \dots, x_n) = 1$ when we choose $x_k = 1$ independently with probability p_k, $k = 1, 2, \dots, n$. For the partial derivatives we have:

$$\frac{\partial \mathbf{P}_{(p_1,\dots,p_n)}(f)}{\partial x_k}|_{(p,p,\dots,p)} = I_k^p(f).$$

Define

$$s_k(f) = \int_0^1 I_k^p(f)dp.$$

$s_k(f)$ is the *Shapley value* of k for f. Clearly,

$$\sum_{k=1}^n s_k(f) = \int_0^1 \sum_{k=1}^n I_k^p(f)dp = \int_0^1 I^p(f)dp = \int_0^1 \frac{d\mathbf{P}(a)}{dp}|dp = \mathbf{P}_1(f) - \mathbf{P}_0(f) = 1.$$

The main results regarding thresholds

Theorem 10.1. *Consider a monotone Boolean function f on n variables. For every $\epsilon, \delta > 0$ there is $\gamma > 0$ such that if the Shapley value $s_i(f) \leq \gamma$ for every $i = 1, 2, \dots n$ then $|I_\epsilon(f)| \leq \delta$. γ can be taken as $K \cdot \log(1/\epsilon)/\log(1/\delta)$, where K is an absolute constant.*

Corollary 10.2. *If f is weakly anonymous then*

$$I_\epsilon(f) \leq K \log(1/\epsilon)/\log n.$$

A more general result is the following:

Theorem 10.3. *For some absolute constant K,*

$$I_\epsilon(f) \leq K \log(1/\epsilon)(\sum_{k=1}^n s_k \log(2/s_k))^{-1}.$$

Remark: These relations may be of some relevance the theory of values of non-atomic games.

10.1 The derivation of the results and related literature

Russo's lemma is a fundamental result in percolation theory (see Grimmett (1989)). The integral representation for the Shapley value is due to Owen (1989).

An early general result concerning threshold behavior is due to Russo (1982). Many of the more recent results rely on a theorem by Kahn, Kalai and Linial (1988) and its extensions. The reader is referred to Talagrand (1994) which contains the sharpest and most general results and to Friedgut and Kalai (1996). Theorems 10.3 can be derived directly from Theorem 1.5 of Talagrand (1984) combined with Owen's representation of the Shapley value.

11 Appendix B: Threshold results for larger alphabets

The known results about threshold properties of Boolean functions are tailor-made for the results in Section 3 when the social choice function give a asymmetric relation. However, some extensions are needed for the more general setting of Section 4.

Let $S = \{s_1, s_2, \dots, s_k\}$ and consider a function $f = f(x_1, x_2, \dots, x_n) : S^n \to S$. Assume that f is monotone in the sense considered in Section 4. Let $\mathbf{p} = (p_1, p_2, \dots p_k)$ be a vector of probabilities, $\sum p_i = 1$. Suppose that each x_i is chosen at random independently and $x_i = s_j$ with probability p_j. We would like to understand the distribution of $f(x_1, x_2, \dots, x_n)$. Define $M_\epsilon(f)$ to be the set of probabilities $\mathbf{p}$ such that the probability that $f = s_k$ is less than $1 - \epsilon$ for every k. $M_\epsilon(f)$ is a subset of the $(k-1)$-dimensional simplex Δ of all probability vectors $\mathbf{p}$ which is the higher dimensional analog of the threshold interval. For our purposes we need theorems that asserts that if f is weakly anonymous and even if the "power" of the individual variables is diminishing then $M_\epsilon(f)$ consists of a "thin membrane" inside the simplex of probabilities.

Theorem 11.1. *Let $S = \{s_1, s_2, \dots, s_k\}$ and consider a neutral, weakly anonymous monotone function $f = f(x_1, x_2, \dots, x_n) : S^n \to S$. The volume of $M_\epsilon(f)$ is bounded from above by a function $H(k, m, \epsilon)$ which, for fixed values of k and ϵ tends to zero as $n \to \infty$.*

The proof of this Theorem will appear elsewhere.

Proof of Theorem 4.1

Let c_0 be a choice function and consider a profile with n_0 individuals such that the plurality leads to c_0. Such a profile exists by Saari's theorem. For an ordering π of the alternatives let $w'(\pi)$ be the number of appearances of the order π and let $w(\pi) = w'(\pi)/n_0$. Consider a random profile on n individuals where for each individual i the probability that i-th preference relation is described by π is $w(\pi)$ (independently for the individuals). Consider now the set U of vectors of probabilities $(p'_s : S \in S)$ where $p'_s \in [p_s + 1/10n_0, p_s + 1/10n_0]$. It follows from Theorem 11.1 that (as n tends to infinity) the value of $f_S(s_1, \dots, s_n)$ is determined with high probability for almost all points in U and from the monotonicity it follows that this value must be t with high probabilities for all points in U. This implies that as n tends to infinity for every S the probability that $c(S) = c_0(S)$ is at least $1 - \delta$. $\square$