Social Choice without Rationality

Gil Kalai *
Institute of Mathematics and Center of Rationality
Hebrew University, Jerusalem, Israel

January 28, 2002

1 Introduction

The purpose of this paper is to discuss two general theorems concerning social choice. Classic social choice theory assumes that the individual choices are "rational" but comes to the conclusion that the social choice cannot always be rational. We will consider more general forms of individual choice and also seek stronger conclusions concerning the society's choice. The two theorems we discuss are from Shelah (2001) and Kalai (2001b). Their proofs are quite involved and will not be presented here. The paper will present the results along with related examples and problems in such a way that will be accessible to a wide audience. For background on social choice theory see Fishburn (1973), Peleg (1984) and Sen (1986). For aggregation in a general framework, see Fishburn and Rubinstein (1986).

Individual choice will be described using choice functions. Thus, given a set X of N alternatives, a *choice function* c is a mapping which assigns to a nonempty subset S of X an element $c(S)$ of S. In decision theory we are mainly concerned with choice functions that are consistent with maximizing behavior. In other words, there is a linear ordering on the alternatives in X and $c(S)$ is the maximum among the elements of S with respect to this ordering. We will refer to such choice functions as *"rationalizable"*.

We will consider general classes of choice functions. These classes are *symmetric*, namely they are invariant under permutations of the elements. Formally, if c is a choice function and π is a permutation from X to X,

*I am thankful to Saharon Shelah, Bezalel Peleg, Noga Alon and Ariel Rubinstein for helpful collaboration

define $\pi(c)(S) = \pi^{-1}c(\pi(S))$. A class $\mathcal{A}$ of choice functions is symmetric if $c \in \mathcal{A}$ implies that $\pi(c) \in \mathcal{A}$ for every permutation π.

We call a class $\mathcal{A}$ of choice functions *proper* if it does not contain all choice functions. When an economic model leads to a proper class of choice functions then there exists a testable implication of the model, i.e., there are choice functions that refute the model. Proper classes of choice functions can also be referred to as classes with *testable implications* (compare with Sprumont (2000)). Of course, the class of rationalizable choice functions is symmetric and has testable implications when there are three alternatives or more.

In this paper we consider general classes of choice functions in the context of social choice theory. In Section 2 we seek an analog of Arrow's impossibility theorem when the individual choices are not necessary rationalizable but belong to a symmetric class of choice function with testable implications. In Section 3 we examine the testable implications of the social choice obtained by the aggregation of a large number of individual rationalizable choices.

Given a society with n individuals, a social choice function is a map $c = F(c_1, c_2, \ldots, c_n)$ which assigns to individual choice functions $c_i \in \mathcal{A}$, $i = 1, 2, \ldots, n$, a choice function c for the society which satisfies the following conditions:

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

(U) (Unanimity) If $c_1(S) = c_2(S) = \cdots = c_n(S)$ then $c(S) = c_1(S)$.

Independence of Irrelevant Alternatives asserts that $c(A)$ can depend only on the individual choice functions within the set A. Independence of Rejected Alternatives is a stronger assumption of a similar nature.

2 Arrow's theorem without rationality

Given a set X of alternatives we will say that a class $\mathcal{A}$ of choice functions defined on non-empty subsets of X satisfies the *Arrow property* if for every non-dictatorial social choice function $F(c_1, \ldots, c_n)$ $(n > 1)$ which satisfies properties (IRA) and (U) and is defined whenever $c_i \in \mathcal{A}, i = 1, 2, \ldots, n$, there is a way to assign a choice function $c_i \in \mathcal{A}$ to every individual i such that for $c = F(c_1, c_2, \ldots, c_n)$, $c \notin \mathcal{A}$.

Conjecture A: Every symmetric class of choice functions with testable implications on at least four alternatives satisfies the Arrow property.

Conjecture A asserts that if there are some testable implication for the class $\mathcal{A}$ of choice functions described by an economic model, then any non-dictatorial mechanism for aggregating individual choices must lead to collective choice behavior which is not in $\mathcal{A}$. In other words, no matter what the testable implications are for the individual choice, if the social choice is not dictatorial some of the implications for the individual choices will be violated for the social choice.

In response to Conjecture A, Shelah (2001) proved a very general result which we now present.

Assuming Independence of Rejected Alternatives, we can restrict our attention to classes of choice functions which are defined for subsets of alternatives of fixed cardinality. Let X be a set of m alternatives and let $P_k(X)$ be the set of subsets of X of cardinality k. A choice function c on $P_k(X)$ is now a function which assigns to every set $S \in P_k(X)$ an element $c(S)$ in S.

Theorem 2.1 (Shelah (2001)). *If $7 \leq k \leq |X| - 7$, for every proper symmetric class $\mathcal{A}$ of choice functions on $P_k(X)$ there does not exist a social choice function that satisfies the following conditions:*

1. *The social choice for S depends only on the individual choices for S.*
2. *If all individuals agree on the choice, then so does the entire society.*
3. *The choice is non-dictatorial.*
4. *The social choice is defined and belongs to $\mathcal{A}$ whenever the individual choices belong to $\mathcal{A}$.*

It would be interesting to replace '7' by '2' in the statement of Shelah's theorem. The case $k = 2$ is especially interesting. There exists a proper class of preference relations ($k = 2$) on three alternatives for which the Arrow property is violated. This is the case when the individual choices on three alternatives are *cyclic*. Namely, if an individual prefers alternative A over B and B over C then it follows that C is preferred over A. This class of preference relations is preserved under majority rule for an arbitrary odd number of individuals.

3 The chaotic nature of social choice

Next we consider rational individuals and ask whether there are any testable implications for the class of choice functions that represent the society's choice.

Conjecture B: When the individual choices are rationalizable, every "genuine" and neutral form of social choice leads for large societies to the full class of choice functions.

By the word "genuine" we (informally) mean that the choice of the society is not determined by the choices of a few individuals. For example, when the society consists of different "types" of individuals, the social choice function does not distinguish between individuals of the same type and there are many individuals of each type. Conjecture B asserts that for a large society every "genuine" social choice function will lead to all choice functions and thus not only will the class of choice functions that arise as the society's choices not be rationalizable but in addition there will be no testable implications for this class.

Conjecture B is supported by a theorem due to McGarvey (1953) which asserts 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. and by a theorem due to Saari(1989) which asserts that the plurality method for sufficiently large societies gives rise to *every* choice function.

A response to Conjecture B using some assumptions on the social choice was given in Kalai (2001b). The meaning of "genuine" was defined in terms of the Shapley value. We will consider the special case of $k = 2$ which involves the aggregation of preference relations. A neutral social choice can be described in terms of a *strong simple game* defined on the set of individuals. A winning coalition is a subset S of individuals such that if every member of S prefers alternative a on alternative b then so does the society. Neutrality also implies that the game is *strong*, namely if S is a winning coalition then the set of individuals not in S is a losing coalition.

Theorem 3.1 (Kalai (2001b)). *For every m, there is $\delta = \delta(m)$ such that if a neutral monotone social choice function is defined on $P_2(X)$, $|X| = m$ is described by a strong simple game G and if the Shapley value of each player in G is at most δ, then the social choice function will lead to all asymmetric preference relations.*

A case in which the Shapley value of each player is small is that in which the individuals are divided into types, such that the social choice is invariant under permutations of individuals of the same type. Another case is when there are "enough" permutations on the individuals for which the simple strong game is invariant. Being invariant to all permutations would leave us with only simple majority games (in which case McGarvey's theorem applies). However, in order to guarantee that all Shapley values are the same it is sufficient that every two individuals be indistinguishable which leads to many additional examples.

Kalai (2001b) also includes general theorems for arbitrary k. For the general case the Independence of Rejected Alternatives is crucial.

Remark: We could combine the two conjectures and ask what happens to "genuine" social choice for many individuals when the individual choices are not rationalizable. There are cases when even for anonymous social choice on an arbitrarily large number of individuals, there are testable implications for the social choice. We have already remarked that the majority rule for individuals with cyclic preference relations on three alternatives yields a cyclic preference relation for the society. More general examples are as follows: When the number of alternatives is $2r + 1$ and the individual preference relations consist of the class of asymmetric relations for which every alternative is preferred over precisely r other alternatives, then when we apply the majority rule with any odd number of individuals there is a simple testable implication for the choice functions that arise: they are *not* rationalizable.

4 Examples

4.1 A specific social choice function

Let $x \in X$ be a fixed alternative and let a and b be two fixed individuals. Consider the following social choice function: $c(S) = c_a(S)$ if $x \in S$ and $c(S) = c_b(S)$ if $x \notin X$. When $\mathcal{A}$ is the class of rationalizable choice functions on $P_k(X)$ then the class of choice functions for the society is a somewhat larger class $\mathcal{A}'$. When the individual choices are from $\mathcal{A}'$ then the social choice will define a larger class $\mathcal{A}^{(2)}$. Define $\mathcal{A}^{(r)}$ in a similar way. Theorem 2.1 implies (if $k \geq 7, |X| - 7$) that $\mathcal{A}^{(r)}$ is the class of all choice functions for some r. In this case, there is a direct proof (for every k).

Let $\mathcal{R}$ denote the class of rationalizable choice functions. The class $\mathcal{R}'$ naturally occurs in certain economic contexts. In this case the choice is according to one order relation for sets S that contain x and according to another for sets S that do not contain x (see, Sen (1993) and Kalai, Rubinstein and Spiegler (2001)). This class is quite limited so that the conclusion of Conjecture B is violated, but so is the principal condition: The social choice does depend on the choices of two individuals.

4.2 Social choice in median land

Consider a society in which the individuals have a "utility function" on alternatives but rather than trying to maximize it they always choose the "median" alternative. Thus, assume that every individual i has a linear ordering on a set X of alternatives. Given a non-empty subset S of odd cardinality of X, the choice $c_i(S)$ of the i-th individual is the middle element of S according to his order. This class of choice functions was considered in Kalai, Rubinstein and Spiegler (2001).

Just like rationalized choice functions based on maximization, such "median" choice functions are based on an ordering of the alternatives, however the two types of choice functions are very different. It is impossible to represent the median as maximization with respect to another ordering (revealed preferences). How would theoretical economics and game theory look in a society where choices are based on the median?

Shelah's theorem shows that Arrow's impossibility theorem can be extended. (when $k \geq 7, |X| - 7$.) It would be interesting to find a simple proof for this example which would also include the cases not covered by Shelah. For the median choice it is not clear whether (IRA) is needed; (IIA) may suffice.

4.3 Social choice for couples

Imagine a society that consists of couples. For a set S of alternatives a couple chooses the most preferred element of S by one of its members. The class of couples' choices has testable implications when the number of alternatives is four or more. (See Kalai, Spiegler and Rubinstein (2001).) Shelah's theorem asserts that there is no non-dictatorial social choice function for this class of choice functions when the number of alternatives is 14 or more.

4.4 An example which demonstrates the connection between the two conjectures

Conjectures A and B are related in the sense that they seem to express similar *mathematical* phenomena [1]: Following is an example which concretely demonstrates this connection:

Consider a non-dictatorial social choice function of the form $c = F(c_1, c_2, c_3)$. We can use F to define a social choice function F_2 on nine individuals as

[1] However, the proofs of the two results are very different.

follows:

$$F_2(c_1, \ldots, c_9) = F(F(c_1, c_2, c_3), F(c_4, c_5, c_6), F(c_7, c_8, c_9)).$$

In a similar way, we can use F to recursively define social choice functions F_k for 3^k individuals as follows:

$$F_k(c_1, \ldots, c_{3^k}) =$$

$$= F(F_{k-1}(c_1, \ldots, c_{3^{k-1}}), F_{k-1}(c_{3^{k-1}+1}, \ldots, c_{2 \cdot 3^{k-1}}), F_{k-1}(c_{2 \cdot 3^{k-1}+1}, \ldots, c_{3^k})).$$

Now, suppose that F enlarges every proper symmetric class of choice functions $\mathcal{A}$ as asserted in Conjecture A. This implies that the class of choice functions result from applying F_{k+1} to the choices of a society of 3^{k+1} rational individuals is strictly larger than the class of choice functions that result from applying F_k to the choices of a society of 3^k rational individuals. It follows that if k is very large, F_k will lead to all choice functions as asserted in Conjecture B.

5 When rejected alternatives count

Assuming only the Independence of Irrelevant Alternatives without assuming the Independence of Rejected Alternatives, the assertions of both Conjectures A and B do not hold. (Of course, when $k = 2$, namely when we consider individual and social preference relations, (IIA) and (IRA) are identical.)

The following example was suggested by Bezalel Peleg: Recall the following way to move from a preference relation to a choice function: 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 vertices of maximal out-degree.

Let $\mathcal{A}$ denote the class of choice correspondences obtained from binary relations R as we have just described. Consider also the class $\mathcal{A}'$ of choice functions obtained from $\mathcal{A}$ by choosing a single element in $c(S)$ according to some fixed order relations on the alternatives.

Now define a social choice function as follows: write aRb if a majority in the society prefers a to b. Note that this rule produces a non-dictatorial social choice and when the individual choices are in $\mathcal{A}$, then so is the social choice and similarly for $\mathcal{A}'$.

In this case the social choice function does not satisfy (IRA): the social choice of S does not depend only on the individual choices for S but also on the preferences among pairs of elements in S. For this social choice function the assertions of both Conjectures A and B are violated.

A similar story can be told about the well-known Borda rule. For every two alternatives a and b, let $w(a,b)$ be the number of individuals i that prefer a on b (namely $c_i(\{a,b\} = a)$. From this list of numbers create a choice correspondence as follows: For $a \in S$ let $u(a,S) \sum\{w(a,s) : s \in S$ and let $c(S)$ be the elements of S for which $u(a,S)$ is maximal. This is the Borda rule and it is well known that when the individual choices are rationalizable it leads to a class $\mathcal{B}$ of choice functions with testable implications. It is proved in Kalai (2001a) that $\log \mathcal{B} \le N^3$. Using the Borda rule (as we have described it), when the individual choices are in $\mathcal{B}$ this yields a choice function for the society which is also in $\mathcal{B}$.

References

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

[2] Fishburn, P. and A. Rubinstein (1986) Algebraic Aggregation Theory, Journal of Economic Theory 38 (1986), 63-77.

[3] Kalai, G., A. Rubinstein and R. Spiegler (2001), Comments on rationalizing choice functions which violate rationality, preprint. `http://www.princeton.edu/~ariel/papers/krs.pdf`

[4] Kalai, G. (2001a) Statistical learnability and rationality of choice, preprint, 2001. `http://www.ma.huji.ac.il/~kalai/L.pdf`

[5] Kalai, G. (2001b) Social choice and threshold phenomena, preprint, 2001. `http://www.ma.huji.ac.il/~kalai/SCT.pdf`

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

[7] Peleg, B. (1984), Game theoretic analysis of voting in committees. Econometric Society Monographs in Pure Theory, 7. Cambridge University Press, Cambridge, 1984.

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

[9] Sen, Amartya (1986) Social choice theory, Ch. 22, Vol III of K. Arrow and D. Intriligator (1986), *Handbook of Mathematical Economics*, North-Holland, Amsterdam, pp. 1073-1182.

[10] Sen, A. (1993) Internal consistency of Choice, Econometrica, 61, 495-521.

[11] Shelah, S., On the Arrow Property, preprint 2001. Available on `http://front.math.ucdavis.edu/math.LO/0112213` (Will be available soon on: `http://www.math.rutgers.edu/~shelah`)

[12] Sprumont, Y. (2000) On the testable implications of collective choice theories, J. Econom. Theory 93, 205–232.