ARROW'S THEOREM

TERENCE TAO

ABSTRACT. Arrow's theorem on voting paradoxes.

1. INTRODUCTION

Suppose we have a finite set V of voters who have to vote for a finite set of candidates C. Each voter ranks the candidates in some ordering, e.g. if $C = \{X, Y, Z\}$ then a voter v might rank $Y > Z > X$. We assume *voter rationality* - that every voter ranks candidates in a total ordering, and *voter free will* - each voter can rank candidates in any order (subject to voter rationality) and independently of all other voters.

A *voting system* is a function which takes as input the voting preferences of each voter in V, and returns as output a ranking of the candidates.

One would like the voting system to satisfy the following reasonable axioms.

- (Voting system rationality) The output of the voting system is a total ordering.
- (Determinism) The output of the voting system is determined only by the ranking preferences of the voters and on no other factors (in particular, no randomization is allowed).
- (Consensus) If all the voters prefer A to B, then the output of the voting system also prefers A to B.
- (Impartiality) All candidates are treated equally; in other words, the voting system is invariant under permutations of the candidates.
- (Independence of a third alternative) The relative ranking of X and Y in the output of a voting system is independent of the voters preferences for a third candidate Z.
- (No dictators) If a single voter prefers X to Y, but all other voters disagree, then the voting system should override the wishes of that single voter and rank Y higher than X.

Unfortunately, we have

Arrow's theorem: If there are at least three candidates, then the above six axioms are inconsistent.

This is a serious obstacle to constructing a "perfect" voting system. One must sacrifice one or more of the axioms in order to create a working voting system. Usually it is the independence of a third alternative which has to go; this is why third parties can "spoil" the results of an election even when they have no realistic chance of winning. A *dictatorship system* (when a single voter dictates the outcome of the voting system) obeys the first five axioms, but this seems an undesirable choice. One amusing choice is a *random dictatorship system*, in which a single voter is chosen by lottery to rank the candidates; this is actually one of the fairest systems,

1991 *Mathematics Subject Classification.* 42B15, 35L05.

but probably impossible to implement in today's political climate. (Plus there is the tremendous practical problem of safeguarding the lottery from corruption).

We remark that the impartiality axiom is not needed for Arrow's theorem, but makes the proof simpler. (The reader is invited to figure out a proof of the theorem without using impartiality - basically one has to deal with separate notions of quorum for each pair of candidates A, B).

A simple example of the problems in a voting system occurs when the number of candidates equals or exceeds the number of voters. Consider for instance the following example of five voters A, B, C, D, E in a group voting for one of their members to be a chairperson:

- A prefers $A > B > C > D > E$.
- B prefers $B > C > D > E > A$.
- C prefers $C > D > E > A > B$.
- D prefers $D > E > A > B > C$.
- E prefers $E > A > B > C > D$.

Note that four out of five voters prefer A to B. Thus, to obey the no dictators axiom, the voting system output must rank A above B. But a similar argument forces the voting system to rank B above C, C above D, D above E, and E above A. But one cannot obey all of these without contradicting voting system rationality.

One may argue that in such cases there should be ties, and that in more realistic situations in which the number of voters far exceeds the number of candidates, this type of cyclic paradox should not occur. For instance when there are only two candidates and an odd number of voters then all seven axioms can be satisfied by a simple majority vote. However, once a third candidate enters one can then leverage the independence axiom to create a contradiction.

The idea is as follows. Define a *quorum* to be any group of voters such that, if all the members in the quorum unanimously vote A above B and all the people not in the quorum vote B above A, then the voting system will place A above B.

The notion of a quorum is well defined; it is not possible for such a group to be able to force a vote some of the time and not at other times because this would contradict the independence axiom. Also, it doesn't matter who A and B are because of the impartiality axiom.

Now for the cyclic part of the argument:

Lemma 1.1. *If S is a quorum, and T is a quorum, then $S \cap T$ is a quorum.*

Proof Choose three candidates A, B, C. Suppose all the voters in S prefer A to B, but none of the voters outside S do. Also suppose all the voters in T prefer B to C, but none of the voters outside T do. Finally suppose all the voters in $S \cap T$ prefer A to C, but none of the voters outside $S \cap T$ do. (This configuration of rankings is possible by voter free will and voter rationality).

Since S is a quorum, the voting system must prefer A over B. Since T is a quorum, the voting system must prefer B over C. By voting system rationality, the voting system must then prefer A over C, which demonstrates that $S \cap T$ is a quorum. ■

The "No dictators" axiom shows that for every voter v, the set $V - \{v\}$ is a quorum (otherwise v would be a dictator). Applying the above lemma repeatedly we thus see that the empty set is a quorum. But this contradicts the consensus axiom, QED.

Voting theory has a vast literature. A sample result (due to Gil Kalai): If there are three candidates, we assume all of the above axioms except for voting system rationality, and all voters vote randomly, then a rational outcome can only be obtained at most 91.92 . . . % of the time. (By comparison, simple majority vote yields a rational outcome 91.22 . . . % of the time, when the number of voters approaches infinity; this seems to be due to Gulibaud). So it seems that voting paradoxes are not isolated to a pathological set of scenarios of infinitesimal probability, but are a genuine and non-trivial difficulty inherent in voting systems.

Department of Mathematics, UCLA, Los Angeles, CA 90024
E-mail address: tao@math.ucla.edu