MS&E 336/CS 366: Computational Social Choice. Winter 2019-20
Course URL: `http://www.stanford.edu/~ashishg/msande336.html`.
Instructor: Ashish Goel, Stanford University.

Lecture 1, 1/8/2020. Scribed by Zhaoyu (Joe) Lou.

1 Introduction and Basic Definitions

This course is about the study of *social choice*, the science of making collective decisions from a computational and algorithmic perspective. We will begin our study with an idealized model for decision making, as follows. We have a set of N **voters**. These voters are choosing between a set of M **candidates**, or **alternatives**. Each voter submits a **ballot** of all M alternatives in the form of a ranking, denoted $\succeq_i$ for voter i. This ranking must satisfy a few properties:

Completeness: For all distinct alternatives a, b, either $a \succeq_i b$ or $b \succeq_i b$.

Transitivity: For any three alternatives a, b, c, if $a \succeq_i b$ and $b \succeq_i c$, then $a \succeq_i c$.

Reflexivity: For any alternative a, $a \succeq_i a$.

Antisymmetry: For two alternatives a, b, if $a \succeq_i b$ and $b \succeq_i a$ then $a = b$. Practically, this simply means that the ranking must be strict.

The sequence of rankings for all voters is called a **profile**. The object of our study in this course is the **social choice function**, a function which maps a profile to a set of winning alternatives. If the output of the social choice function is always a singleton set, it is called **resolute**. We additionally define a **social welfare function**, which instead outputs a complete ranking (not necessarily strict) of the alternatives, although this is sometimes called a social choice function as well.

2 Axioms of Social Choice

We approach the problem of defining "good" social choice functions from an axiomatic perspective. There are many intuitive and natural desiderata for social choice functions, including:

Pareto Optimality: For alternatives a, b, if $a \succeq_i b$ for all voters i, then b is not in the winning set (for a social choice function) or $a \succeq b$ in the output ranking (for a social welfare function).

Non-Imposition: A weaker condition than Pareto optimality, this simply requires that for all alternatives a, there exists a profile of ballots such that a is the winner output by the social choice function.

Anonymity: The output of the social choice/welfare function is invariant to permutations in the labels of the voters.

Neutrality: Permuting alternatives in the ballots likewise permutes them in the output of the social choice/welfare function.

Reinforcement: Let S and T be two profiles, and let f be our social choice function. If $f(S) \cap f(T) \neq \varnothing$, then $f(S+T) \subseteq f(S) \cap f(T)$.

Strategy Proofness: Informally, voters have no incentive to vote anything other than their true preferences. This will be defined more formally in our discussion of game theory in the context of social choice. We will see that this axiom in particular presents a challenge for designing social choice functions.

Independence of Irrelevant Alternatives (IIA): The relative order of two alternatives a and b in the output of the social choice function depends only on the relative order of a and b in each voter's ranking.

Our goal will be to find social choice functions which satisfy as many of these axioms as possible (ideally all of them). Unfortunately, we will see this is more challenging than it first appears.

3 Social Choice Functions

The field of social choice started over two centuries ago, and social choice functions have largely fallen into two fundamentally different camps ever since. One camp favors a combinatorial approach, using as a litmus test the **Condorcet criterion**, which requires that if an alternative beats all other alternatives in pairwise elections according to the rankings submitted, then it wins. (Note that it is possible to have a **Condorcet paradox** where no Condorcet winner exists. A simple example is a three voter, three candidate case where one voter submits a ranking $a \succeq_1 b \succeq_1 c$, one submits $b \succeq_2 c \succeq_2 a$, and one submits $c \succeq_3 a \succeq_3 b$.)

The Condorcet criterion admits a very natural extension to a social choice function, the **Copeland rule**. In the Copeland rule, the winning set is the set of alternatives which win the most pairwise elections. The Copeland rule is Pareto optimal, anonymous, neutral, and respects IIA, but it is not reinforcing. It is strategy proof when restricted to profiles which have a Condorcet winner.

The other camp of social choice functions favors a scoring based approach, where alternatives are assigned score values based on the ballots and the scores define the output. The most natural of these is the **Borda rule**, in which for each ballot, the top ranked alternative recieves $M-1$ points, the second recieves $M-2$, and so on down to 0 points for the lowest ranked alternative. The alternative(s) with the highest total points win. The Borda rule is Pareto optimal, anonymous, reinforcing, and neutral, but does not respect the Condorcet criterion or IIA. This scoring scheme is also equivalent to assigning each alternative a number of points equal to the number of ballots on which it beats another alternative, summed over all other alternatives.

We can also consider other score assignment schemes, which give rise to a family of **scoring rules**. One example is the **plurality** scheme, which simply says that the alternative with the most first place votes wins. This is equivalent to assigning a score of 1 to the top ranked alternative and 0 to all others. (Note that this differs from the **majority** scheme, which requires the winner to have a majority of top ranked votes.)

4 Impossibility Results

A major impetus behind the existence of the field of social choice is the unfortunate fact that it turns out that many of the axioms we enumerated as desiderata are mutually exclusive. In particular, it turns out that simultaneously satisfying some of the axioms creates a **dictatorial** system in which a single privileged voter determines the outcome, a clear (and especially egregious) violation of the axiom of anonymity. The most famous of these impossibility theorems is **Arrow's Theorem**.

Arrow's Theorem: Any social welfare function which is Pareto optimal and satisfies IIA is dictatorial.

However, in our class we will be more concerned about a response to an even stronger impossibility result which drops the somewhat unnatural IIA axiom for a much weaker, more natural criterion. This is the **Gibbard-Satterthwaite Theorem**.

Gibbard-Satterthwaite Theorem: Any resolute, strategy-proof social choice function which is non-imposing for at least 3 alternatives is dictatorial.

This is a very strong impossibility result, and this class is largely about responding to this theorem. What relaxations of our axioms or restrictions of our domain can we make to get reasonable guarantees on a social choice function?