MS&E 336/CS 366: Computational Social Choice. Win 2023-24
Course URL: `http://www.stanford.edu/~ashishg/msande336/index.html`.
Instructor: Ashish Goel, Stanford University.

Lecture 1, 1/16/2024. Scribed by Jiaju Liu.

Arrow's Impossibility Theorem

Arrow's Impossibility Theorem is one of the first impossibility proofs covered in this course. Before stating it, we will introduce two prerequisite voting desiderata.

Pareto Optimality

For two alternatives a and b, if $a \prec_i b$ for all voters i, then a must strictly be ranked higher than b by the social welfare function (SWF) and be cannot be the output of the social choice function (SCF). This means that an alternative cannot win an election if every voter strictly prefers some other alternative.

Independence of Irrelevant Alternatives

The Independence of Irrelevant Alternatives (IIA) criteria states that the relative ordering of two alternatives given a social welfare function must be solely defined by relative voter opinions on those two alternatives. Specifically, changing voter preferences of alternatives c, d, e and so on should not alter the relative ranking of alternatives a and b.

Another framing of IIA is as follows: if a set of voters agree on some alternative x and we remove y from the space of alternatives, then the voters should still agree on alternative x. Here, y is an irrelevant alternative.

Theorem Statement

After defining these two desiderata, we can now introduce Arrow's Impossibility Theorem.

Theorem 2.1: Any SWF among three or more alternatives that is Pareto Optimal and satisfies IIA must be dictatorial.

By dictatorial, we mean that the outcome of the vote is determined solely by one voter's preferences. This theorem says that, only requiring a criterion as weak as Pareto Optimality, no interesting voting rules can satisfy IIA. No extra desiderata need to be added for this impossibility result; it applies regardless of anonymity, neutrality, strategy proofness, etc. However, this result is not too discouraging, as IIA is a strong requirement, and in future sections we will talk about different voting rules that accomplish other goals.

Voting Rules

Scoring Rules

Scoring rules work by assigning weights to each position in a ranking. If there are M candidates, assign every ranking from 1 to M a score according to the score vector $\mathbf{r} = (r_1, r_2, \ldots, r_M)$. The score vector r is proper if $r_1 \geq r_2 \geq ...r_M \geq 0$. Given a profile, each candidate's score is the sum of the weights across all of the rankings they receive. For example, look at the following profile with 4 voters, 3 candidates, and score vector $\mathbf{r} = (2, 1, 0)$:

Score	**1**	**2**	**3**	**4**
2	a	a	c	a
1	c	b	a	c
0	b	c	b	b

Then candidate a's score is $2 + 2 + 1 + 2 = 7$, candidate b's score is $0 + 1 + 0 + 0 = 1$ and candidate c's score is $1 + 0 + 2 + 0 = 3$,. The SWF ranks candidates by decreasing score (a,c,b) and SCF chooses the candidate with the highest score (a). Some special cases of scoring rules are:

Borda: Choose $r_1 = M - 1, r_2 = M - 2, \ldots, r_M = 0$. Notice that r_i corresponds to the number of candidates that a candidate ranked in the ith position is beating. The previous example uses Borda scoring.

Plurality: Choose $r_1 = 1$ and $r_2 = r_3 = \cdots = r_M = 0$. This is a common scoring rule we see in elections.

Scoring rules are always anonymous, neutral, reinforcing, and Pareto optimal if the scores are strictly decreasing. However, they do not satisfy the Condorcet criterion and are not strategy-proof. We will use the Borda rule as an example to illustrate non-strategy proofness, but note that no scoring rule is strategy proof.

Consider the following profile

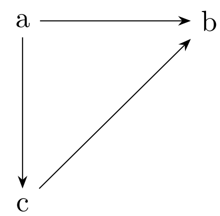
Score	**1**	**2**
3	a	c
2	b	b
1	d	a
0	c	d

Currently candidates a and b are tied for winning. However, voter 2 can make sure that candidate b wins by misreporting candidate b as their top choice.

Copeland Scoring

In the previous example, candidate c beats b in a pairwise election. We can see this because removing candidate a from the profile results in the modified profile where c wins:

1	**2**	**3**	**4**
c	b	c	c
b	c	b	b



The Copeland rule scores candidates by the number of alternatives they beat in a pairwise election. Specifically, in an election over a set of candidates A, the score of candidate $x \in A$ is

$$Copeland(x) = |\{x' \in A | \texttt{x beats x' pairwise}\}| - |\{x' \in A | \texttt{x' beats x pairwise}\}|.$$

Since candidate c beats b but loses to a in their respective pairwise elections, c's Copeland score is $Copeland(c) = 1 - 1 = 0$. Another way to visualize the results of pairwise election is to build a tournament graph. Tournament graphs are directed graphs on the set of candidates where there is an edge from x to x' if x beats x' in a pairwise election. The tournament graph for the example profile is then:

A candidate's Copeland score is the number of edges directed away from the node minus the number of edges directed to the node. Finally, the Copeland winner is the candidate with the highest Copeland score. Since $Copeland(a) = 2$ and $Copeland(b) = -2$, a is the Copeland winner.

The Copeland rule is anonymous, neutral, Pareto optimal, and satisfies the Condorcet criterion. However, it is not reinforcing. In addition, this rule is not strategy-proof and results can be manipulated in unintuitive ways. Consider the following profile across 5 candidates and 7 voters:

1	**2**	**3**	**4**	**5**	**6**	**7**
a	a	d	d	d	e	e
b	b	e	e	e	c	c
c	c	b	b	b	a	a
d	d	c	c	c	d	d
e	e	a	a	a	b	b

The Copeland scores across candidates are

$$Copeland(a) = 2 - 2 = 0$$
$$Copeland(b) = 1 - 3 = -2$$
$$Copeland(c) = 2 - 2 = 0$$
$$Copeland(d) = 2 - 2 = 0$$
$$Copeland(e) = 3 - 1 = 2$$

Given this profile, candidate e wins. However, voter 1 could change the election to be more preferable to them by swapping the position of candidates a and d in their profile. This would increase candidate d's Copeland score to 4, making it the new winner. Therefore, voter 1 can achieve a better outcome by misreporting their preferences.

Approval Voting

Approval voting assumes that voters have dichotomous preferences. Specifically, among a set of candidates A each voter v has a preferred set $C_v \subseteq A$ such that

$$Utility_v(c) = \begin{cases} 1 & c \in C_v \\ 0 & c \notin C_v \end{cases}$$

We call C_v the approval set for voter v. This assumes that the voter likes all of their approved candidates equally and dislikes everyone else equally. The winner with approval voting is the most candidate that is in the most number of people's approval sets. Approval voting is both welfare maximizing and strategy proof!