22 Social Choice

Social choice theory asks how the possibly conflicting preferences of a set of agents can be aggregated into a collective decision, and in particular which properties the aggregate choice should satisfy and which properties can be satisfied simultaneously. Examples of settings that can be studied in the framework of social choice theory include voting, resource allocation, coalition formation, and matching.

22.1 Social Welfare Functions

Let $N = \{1, \ldots, n\}$ be a set of agents, or *voters*, and $A = \{1, \ldots, m\}$ a set of alternatives. Assume that $n, m \geqslant 2$ and finite. Assume that each voter $i \in N$ has a strict linear order $\succ_i \in L(A)$ over A, and the goal is to map the profile $(\succ_i)_{i \in N}$ of individual preference orders to a social preference order. This is achieved by means of a *social welfare function* (SWF) $f : L(A)^n \to L(A)$.

When $m = 2$, selecting the social preference order that is preferred by a majority of the voters is optimal in a rather strong sense. An SWF $f : L(A)^n \to L(A)$ is

- *anonymous* if for every permutation $\pi \in S_n$ of the voters and all preference profiles $\succ, \succ' \in L(A)^n$ such that $a \succ_i b$ if and only if $a \succ'_{\pi(i)} b$ for all $a, b \in A$, it holds that $f(\succ) = f(\succ')$;
- *neutral* if for every permutation $\pi \in S_m$ of the alternatives and all preference profiles $\succ, \succ' \in L(A)^n$ such that $a \succ_i b$ if and only if $\pi(a) \succ'_i \pi(b)$ for all $a, b \in A$, it holds that $a \; f(\succ) \; b$ if and only if $\pi(a) \; f(\succ') \; \pi(b)$ for all $a, b \in A$; and
- *monotone* if for all $\succ, \succ' \in L(A)^n$ and $a, b \in A$, $a \; f(\succ) \; b$ and $\{i \in N : a \succ_i b\} \subseteq \{i \in N : a \succ'_i b\}$ imply that $a \; f(\succ') \; b$.

Anonymity requires that voters are treated equally, symmetry that alternatives are treated equally, and monotonicity that an alternative cannot become less preferred socially when it becomes more preferred by individuals. When the number of voters is odd, these intuitive fairness and welfare properties precisely characterize the majority rule.

Theorem 22.1. *Consider an SWF $f : L(A)^n \to L(A)$, where $|A| = 2$ and n is odd. Then f is the majority rule if and only if it is anonymous, neutral, and monotone.*

Proof sketch. Let $A = \{a, b\}$. By anonymity, the social preference only depends on the number of voters that prefer a to b. By neutrality, the social preference has to change between a preference profile where $\lfloor n/2 \rfloor$ voters prefer a to b and one where $\lceil n/2 \rceil$ voters prefer a to b. By monotonicity, the socially preferred alternative can never change from a to b when the number of voters who prefer a to b increases, so this is actually the unique change, and it follows that f is the majority rule. $\square$

a	b	c
b	c	a
c	a	b

Figure 22.1: An instance of the Condorcet paradox. Each column lists the preferences of a particular voter.

The result also holds for weak preference orders if monotonicity is replaced by positive responsiveness, which requires that a weak social preference for b over a changes to a strict preference for a when some voter changes from a strict preference for b to a weak preference for a or from a weak preference for b to a strict preference for a.

In light of this result, it seems promising to base the decision on pairwise comparisons of alternatives even when $m > 2$. As the Marquis de Condorcet pointed out already in 1785, this is somewhat problematic, since the pairwise majority relation may contain cycles. To see this, consider a situation with three alternatives a, b, and c, and three voters with preferences as shown in Figure 22.1. It is easily verified that a majority of the voters prefers a over b, a majority prefers b over c, and a majority prefers c over a.

Unfortunately, this kind problem is not specific to the majority rule, but applies to every SWF satisfying a set of desirable criteria. An SWF $f : L(A)^{|N|} \to L(A)$ is

- *Pareto optimal* if for all $a, b \in A$ and every $\succ \in L(A)^n$ such that $a \succ_i b$ for all $i \in N$, it holds that $a \succ' b$, where $\succ' = f(\succ)$;
- *independent of irrelevant alternatives* (IIA) if for all $a, b \in A$ and all $\succ, \succ' \in L(A)^n$ such that $\succ_i \cap (\{a,b\} \times \{a,b\}) = \succ'_i \cap (\{a,b\} \times \{a,b\})$ for all $i \in N$, it holds that $f(\succ) \cap (\{a,b\} \times \{a,b\}) = f(\succ') \cap (\{a,b\} \times \{a,b\})$; and
- *dictatorial* if there exists $i \in N$ such that for all $\succ \in L(A)^n$, $f(\succ) = \succ_i$.

Pareto optimality requires that alternative a is socially preferred over alternative b when every voter prefers a over b. Independence of irrelevant alternatives requires that the social preference with respect to a and b only depends on individual preferences with respect to a and b, but not on those with respect to other alternatives. Finally, an SWF is dictatorial if the social preference order is determined by a single voter. It turns out that dictatorships are the only SWFs for three or more alternatives that are Pareto optimal and IIA.

Theorem 22.2 (Arrow, 1951). *Consider an SWF $f : L(A)^n \to L(A)$, where $|A| \geqslant 3$. If f is Pareto optimal and IIA, then f is dictatorial.*

Requiring non-dictatorship and Pareto optimality is rather uncontroversial. Relaxing IIA for example enables *Kemeny's rule*, which chooses a social preference order $\succ'$ that maximizes the number of agreements with the individual preferences, such that

$$\sum_{i \in N} |\succ' \cap \succ_i| = \max_{\succ'' \in L(A)} \sum_{i \in N} |\succ'' \cap \succ_i|. \tag{22.1}$$

4	3	2
a	c	b
b	b	c
c	a	a

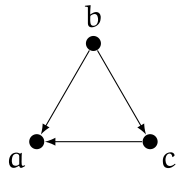

Figure 22.2: Preferences of three types of voters over a set of three alternatives (left), and the graph of the corresponding majority relation (right). Each column of the table on the left lists the preferences of a particular type of voter, the number of voters of that type is given in the top row. In the graph on the right, a directed edge from alternative x to alternative y indicates that a majority of the voters prefer x to y.

This maximization problem is NP-hard, but can be written as an integer program. An interesting alternative characterization of Kemeny's rule is as the *maximum likelihood estimator* for a simple probabilistic model in which votes are generated by an underlying "true" preference order. Fix a preference profile $\succ \in L(A)^n$ and define a vector $w \in \mathbb{N}^{m \times m}$ by letting w_{xy} be the number of voters who prefer x to y, i.e., $w_{xy} = \{i \in N : x \succ_i y\}$. Now assume that this vector was instead generated by picking a single preference order $\succ' \in L(A)$ and a probability $p \in (1/2, 1]$, and for each voter $i \in N$ and each pair of alternatives $x, y \in A$, letting i choose between x and y according to $\succ'$ with probability p and opposite to $\succ'$ with probability $1-p$. The probability of obtaining vector w from preference order $\succ'$ would then be

$$\mathbb{P}(w \mid \succ') = p^{\sum_{i \in N} |\succ_i \cap \succ'|}(1-p)^{nm(m-1)/2 - \sum_{i \in N} |\succ_i \cap \succ'|},$$

and it is easy to see that this probability is maximized by a preference order $\succ'$ that satisfies (22.1).

22.2 Social Choice Functions

Instead of the explicit assumptions of Theorem 22.2, one could also relax an implicit one. In particular, it might often suffice to identify a single best alternative rather than giving a complete ranking. This is achieved by a *social choice function* (SCF) $f : L(A)^n \to A$. Two of the most familiar SCFs are *plurality*, which chooses an alternatives ranked first by the largest number of voters, and *single transferable vote* (STV), which successively eliminates alternatives ranked first by the fewest voters until only one alternative remains.

Consider for example a situation with three alternatives a, b, and c, and nine voters with preferences as shown on the left of Figure 22.2. In this situation, plurality selects alternative a because it is ranked first by 4 voters, compared to 3 for c and 2 for b. STV first eliminates alternative b, which is ranked first by only 2 voters. Restricting attention to the remaining alternatives, a is ranked first by 4 voters and c by 5 voters. Alternative a is thus eliminated next, while alternative c remains and is selected. The graph of the majority relation shown on the right of Figure 22.2 illustrates that alternative b is a so-called *Condorcet winner*, i.e., it is preferred to any other alternative by a majority of the voters, while alternative a is a *Condorcet loser*, i.e., a majority of voters prefer any other alternative to a. The example of Figure 22.1 shows that a Condorcet winner or loser need not exist, but it is certainly reasonable to require that a Condorcet winner is selected when it exists, and that a Condorcet loser is never selected. An SCF satisfying the former property is called Condorcet consistent, and the example of Figure 22.2 shows that that neither plurality nor STV are Condorcet consistent.