MS&E 336/CS 366: Computational Social Choice. Aut 2010-21
Course URL: `http://www.stanford.edu/~ashishg/msande336/index.html`.
Instructor: Ashish Goel, Stanford University.

Lecture 2, 9/22/2021. Scribed by Philippe Pangestu and Roberto Saitto.

Notation

Let $I = \{1, ..., N\}$ and $A = \{1, ..., M\}$ be the set of individuals and the set of alternatives, respectively. Let $\gtrsim \equiv (\gtrsim_i)_{i \in I}$ denote a general profile of individual orderings and let U be the set of all such profiles. Following Sen (2017), the term ordering refers to a complete, reflexive and transitive binary relation.

Desiderata for Collective Decisions, continued

Independence of irrelevant alternatives

A very famous and important property which one may impose on a social welfare function is based on the principle that the social preference over some pair of alternatives $x, x' \in A$ should exclusively depend of the individual preferences over x and x'. That is, it is argued that the individual preferences over some pair of alternatives y, y' should be irrelevant in determining the social preference over any other distinct pair of alternatives x,x'. Consistently, this property is called independence of irrelevant alternatives (IIA).

Definition 2.1. *A social welfare function f is independent of irrelevant alternatives if for all pairs of profiles of individual orderings $\gtrsim, \gtrsim'$ and all alternatives $x, x' \in A$ such that for all individuals $i \in I$*

$$x \gtrsim_i x' \textit{ if and only if } x \gtrsim'_i x$$

being $R = f(\gtrsim)$ and $R' = f(\gtrsim')$ one has

$$xRx' \textit{ if and only if } xR'x'.$$

An equivalent restriction may be imposed on social choice functions, observing that any SCF can be induced by a mapping from the set of individual orderings to the set binary relations over A. Indeed, being $R(\cdot)$ such a map, one can define

$$f(\gtrsim) = \left\{x \in A \mid xR(\gtrsim)x' \text{ for all } x' \in A\right\}.$$

Note that in lecture 1 a SCF has been defined as a map between the set of individual orderings and the power set of the set of alternatives, $\mathscr{P}(A)$. Under this definition, each map $R(\cdot)$ induces a unique SCF, but a given SCF f can be induced by any map $R(\cdot)$ such that for all all profiles $\gtrsim$ and $x, x' \in A$

$$xR(\gtrsim)x' \text{ if } \ x \in f(\gtrsim) \ \text{ and}$$

$$\text{if } x \notin f(\gtrsim), \text{ then } \exists x' \in A \text{ s.t. } \neg\left[xR(\gtrsim)x'\right].$$

Also observe that a SCF induced by a map $R(\cdot)$ according to the above construction may be empty-valued for some profile of individual orderings. Indeed, one can show that a map $R(\cdot)$ induces a non-empty valued SCF if and only if its image is contained in the set of complete, reflexive and acyclical[1] binary relations over A.

This construction sheds light on the fact that a SCF is significantly less restricted than a SWF. Indeed, recall that a SWF is a map between the sets of individual orderings to the set of orderings over A. [2]. Intuitively, this reflects the fact that a SWF aims to rank all the alternatives, while a SCF only aims to pick an element from the alternative set.

With this in mind, one may give the following definition for a social choice function to be independent of the irrelevant alternatives.

Definition 2.2. *A social choice function f is independent of irrelevant alternatives if it can be induced by a map $R(\cdot)$ from the set of individual orderings to the set of complete, reflexive and acyclical binary relations over A which satisfies the following. Let $\gtrsim, \gtrsim'$ be a pair of individual ordering profiles such that for some $x, x' \in A$, one has*

$$x \gtrsim_i x' \textit{ if and only if } x \gtrsim'_i x'$$

for all individuals $i \in I$. Then,

$$xR(\gtrsim)x' \textit{ if and only if } \ xR\left(\gtrsim'\right)x'.$$

The argument in support of independence of irrelevant alternatives under its social choice function declination is particular intuitive. Indeed, the requirement can be informally expressed as it follows. Suppose that for a given profile of individual orderings the social choice function selects $x \in A$. Then, it seems natural for x to be chosen even if some individual preferences over some other alternatives x', x'' is modified.

[1]Let R be a binary relation over A and let P be its strict part, that is let xPx' if and only if xRx' and $\neg(x'Rx)$. A binary relation R is acyclical if for all $x_1, ..., x_n \in A$ such that $x_k P x_{k+1}$ for all $k \in \{1, ..., n-1\}$ one has $x_1 R x_n$.

[2]Recall from lecture 1 that an ordering is a complete, reflexive and transitive binary relation and observe that transitivity implies acyclicity.

Condorcet consistency

A desirable property one may require to be satisfied by a SCF is to select a Condorcet winner whenever possible. Recall that a Condorcet winner is an alternative $x \in A$ which is weakly preferred to all other alternatives $x' \in A$ by at least half of the individuals, and that it may not exists for some profile of individual orderings. A SCF which obeys this requirement is said to satisfy the Condorcet criterion, or to be Condorcet consistent. Formally, one can give the following definition.

Definition 2.3. *A social choice function f is Condorcet consistent if for all alternatives $x \in A$, whenever $\gtrsim$ is a profile of preference orderings such that x is a Condorcet-winner then $x \in f(\gtrsim)$.*

Arrow's General Impossibility Theorem

The most famous result in social choice theory is due to the economist Kenneth J. Arrow, and it is a negative result. It shows that, whenever there are at least three alternatives, there is no non-dictorial SWF which is independent of irrelevant alternatives and satisfies a weak form of Pareto optimality, stated in the following definition.

Definition 2.4. *A social welfare function f is weakly Pareto optimal if for all $x, x' \in A$, whenever the profile of individual orderings $\gtrsim$ is such that $x >_i x'$ for all individuals $i \in I$, being $R = f(\gtrsim)$ one has xRx' and $\neg(x'Rx)$.*

Observe that weak Pareto optimality is always implied by (strong) Pareto optimality, which only requires $x \gtrsim_i x'$ for all $i \in I$ and $x >_j x'$ for some $j \in I$ to have xRx' and $\neg(x'Rx)$. Then, one can prove the following result, known as the general impossibility theorem.

Theorem 2.1. *If the set of alternatives has cardinality greater than 2, any social welfare function which is weakly Pareto optimal and independent of irrelevant alternatives is dictorial.*

Proof: Omitted, see the course website for more details. □

It is interesting to remark that it has been shown that the result collapses when considering social choice functions. That is, expanding the set of admissible maps from all the social welfare functions to all the maps which have image within the set of complete, reflexive but only acyclical binary relations is sufficient to escape Arrow's impossibility result. That is, Arrow's argument crucially relies on the transitivity of the social orderings, which may be considered an unnecessary restriction when one only aims to make a choice among all the alternatives. However, since Arrow's theorem, a extremely large number of similar other impossibility results has been proved for even more general settings.

Voting Rules

General voting rules

A voting rule, or scoring rule, assigns a weight to each position in the ranking. Each alternative is assigned a score equal to the sum of its score in all the individual rankings. A SWF (SCF) follows a voting rule if it ranks (chooses) the alternatives in order of total weights. Assume that individual orderings are strict. To formally define a voting rule, it is necessary to rank the alternatives according to each individual order. To this end, for all $i \in I$, let $A_0^i(\gtrsim) = A$ and for all $k \in \{1, ..., M\}$ recursively define

$$x_k^i(\gtrsim) = \left\{x \in A_{k-1}^i(\gtrsim) \mid x \gtrsim_i x' \text{ for all } x' \in A\right\}$$

$$A_k^i(\gtrsim) = A_{k-1}^i(\gtrsim) \setminus x_k^i(\gtrsim).$$

Note that $x_k^i(\gtrsim)$ is always a singleton since the orderings are strict. Furthermore, for all individuals $i \in I$, the family $\left\{x_k^i(\gtrsim)\right\}_{k \in A}$ is a partition of A. Next, one may define a general score function as it follows.

Definition 2.5. *A scoring function $v : (A \times U) \times \mathbb{R} \to \mathbb{R}$ is defined as it follows. For all $w \in \mathbb{R}^M$, alternatives $x \in A$ and preference profiles $\gtrsim \in U$, define*

$$v^i((x, \gtrsim), w) = \sum_{k \in A} 1_{x_k^i(\gtrsim)}(x)\, w_k$$

where $1_{x_k^i(\gtrsim)} : A \to \{0, 1\}$ is such that

$$1_{x_k^i(\gtrsim)}(x) = \begin{cases} 1 & \textit{if } x \in x_k^i(\gtrsim) \\ 0 & \textit{otherwise} \end{cases}$$

for each individual $i \in I$. Then,

$$v((x, \gtrsim), w) = \sum_{i \in I} v^i(x, \gtrsim; w).$$

Finally, it is possible to give a formal definition of voting rule and establish its connections with SWFs and SCFs.

Definition 2.6. *A voting rule is a vector $w \in \mathbb{R}^M$. A social welfare function f follows a voting rule w if for all profiles $\gtrsim$, being $R = f(\gtrsim)$, for all $x, x' \in A$ one has*

$$xRx' \textit{ if and only if } v((x, \gtrsim), w) \geq v\left((x', \gtrsim), w\right).$$

A social choice function f follows a voting rule w if

$$f(\gtrsim) = \left\{ x \in A \mid v(x, \gtrsim, w) \geq v\left(x', \gtrsim, w\right) \text{ for all } x' \in A \right\}$$

for all profiles $\gtrsim$.

Observe that a voting rule can always define a SWF since the scoring function always generate an ordering over A. For all voting rules w, one may say that w satisfies a given property, e.g. Pareto optimality, if its associated SWF or SCF satisfies that property.

Examples and properties of voting rules

As it is clear from the definition, there is an (uncountable) infinity of voting rules. The plurality rule and the Borda count are among the most famous. The plurality rule assigns a score of 1 the candidate who is ranked first and zero to all the others. This voting rule is used in many countries for their elections. Indeed, all first-past-the-post voting systems, like the UK general elections, are consistent with the plurality rule.

Definition 2.7. *The plurality rule is a voting rule* $w \in \mathbb{R}^M$ *such that for all* $k \in \{1, ..., M\}$ *one has* $w_k = 1$ *if* $k = 1$ *and* $w_k = 0$ *if* $k > 0$.

On the other hand, despite the fact that it is quite rare to observe the Borda count in the real world, this voting rule is theoretically relevant and has received great attention from the literature. The Borda count assigns a score of $M - k$ to the $k-$th ranked alternative. An intuitive explanation of this method is given by the following observation: the score of each alternative is exactly the number of alternatives to which it is preferred.

Definition 2.8. *The Borda count is a voting rule* $w \in \mathbb{R}^M$ *such that for all* $k \in \{1, ..., M\}$ *one has* $w_k = M - k$.

A nice property of the Borda count is that evaluating the Borda score on a random sample of individuals provides an unbiased estimate of the true Borda score. The following is an example to gain confidence with the Borda count and the plurality rule. It also shows that the Borda count is different from the plurality rule.

Example 1. *Let* $I = \{1, 2, 3, 4, 5\}$ *and* $A = \{a, b, c, d\}$ *and consider the following profile* $\gtrsim$ *of* *individual orderings.*

	Individual i $\rightarrow$	1	2	3	4	5
Borda w $\downarrow$	*Plurarility rule* w' $\downarrow$					
3	1	a	a	d	d	b
2	0	b	c	b	b	a
1	0	c	b	a	c	c
0	0	d	d	c	a	d

Then, the Borda scores are $v((a, \gtrsim), w) = 9$*,* $v((b, \gtrsim), w) = 10$*,* $v((c, \gtrsim), w) = 5$ *and* $v((d, \gtrsim), w) = 6$*. On the other hand, the plurality scores are* $v((a, \gtrsim), w') = 2$*,* $v((b, \gtrsim), w') = 1$*,* $v((c, \gtrsim), w') = 0$ *and* $v((d, \gtrsim), w') = 2$*. Thus, it follows that the Borda count is different from the plurality rule.*

A very natural restriction on voting rules consists in imposing that a the scoring function must be increasing in the ranking over A defined by the ordering of each individual. Clearly, this property is satisfied by both the plurality rule and the Borda count.

Definition 2.9. *A voting rule is decreasing if for all* $k \in \{1, ..., M-1\}$ *one has* $w_k > w_{k+1}$*.*

Voting rules satisfies many desirable properties. All voting rules are anonymous (and thus non-dictorial) and neutral. Properly adapting the domain, they are reinforcing. Finally, one can see that all decreasing voting rules are Pareto optimal. However, they have also many shortcomings. First, recalling that the scoring function always generates an ordering over A, given that they are non-dictorial and Pareto optimal, as a corollary of Arrow's theorem it follows that no decreasing voting rule is independent of irrelevant alternatives. Moreover, no voting rule is Condorcet consistent and no voting rule is strategy-proof. What follows focuses on Condorcet consistency. The next example shows that the Borda count is not Condorcet consistent.

Example 2. *Let* $I = \{1, 2, 3\}$ *and* $A = \{a, b, c, d, e\}$ *and consider the following profile* $\gtrsim$ *of individual orderings.*

Individual i $\rightarrow$	1	2	3
Borda count w $\downarrow$			
4	a	a	b
3	b	b	e
2	c	c	d
1	d	d	c
0	e	e	a

Then, a the preferences of 1 *and* 2 *ensure that* a *is a Condorcet winner. However, the Borda scores are such that* $v((a,\gtrsim),w)=8$ *while* $v((b,\gtrsim),w)=10$*. Thus, the Borda count is not Condorcet consistent.*

The next example shows that the plurality rule is not Condorcet consistent.

Example 3. *Let* $I=\{1,2,3,4,5\}$ *and* $A=\{a,b,c,d\}$ *and consider the following profile* $\gtrsim$ *of individual orderings.*

Individual i →	1	2	3	4	5
Plurality rule w ↓					
1	a	b	b	d	c
0	b	a	a	a	a
0	c	c	c	b	b
0	d	d	d	d	d

Then, a *is preferred to* b *by* 3 *individuals, and to* c *and* d *by* 4 *individuals. Thus,* a *is a Condorcet winner. However, the plurality scores are such that* $v((a,\gtrsim),w)=1$ *while* $v((b,\gtrsim),w)=2$*. Thus, the plurality rule is not Condorcet consistent.*

In general one can state the following proposition.

Proposition 2.2. *There is no Condorcet consistent voting rule.*

Proof: Omitted, it will be part of an homework. □

Copeland Social Choice Function

Let $a >^{\mu} b$ when a is ranked higher than b in the majority of the rankings in the profile. We then Define the Copeland score, $Copeland(x)$ as[3]

$$Copeland(x)=|\{y\in A \mid x>^{\mu} y\}|-|\{y\in A \mid x<^{\mu} y\}|$$

In other words, the Copeland score of x is calculated by the number of alternatives that x beats in a pairwise election minus the number of alternatives that it loses in a pairwise election. In a way it is similar to the Borda score defined in the previous section, except this time we ignore the margin of victory (by how much x wins) between two alternatives in a pairwise election. The Copeland social choice function selects the set of winners to be

[3]Note that in class, Prof. Goel defined the Copeland rule differently. It was defined as $Copeland(x)=|\{y\in A \mid x>^{\mu} y\}|$ instead. It can be proven that the two definitions are equivalent as the two SCFs always output the same outcomes.

the alternatives with the highest Copeland score. It is easy to see that the Copeland SCF is Condorcet consistent by definition.

The Copeland score is anonymous, neutral, pareto optimal and not reinforcing. The Copeland SCF seems good in terms of Condorcet consistency, however it does have its vulnerabilities. Consider the following example.

Example 4. *Let* $I = \{1, 2, 3\}$ *and* $A = \{a, b, c, d\}$ *and consider the following profile* $\gtrsim$ *of individual orderings.*

Individual i →	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

Then, we see that the Copeland scores for the alternatives a, b, c, d *are respectfully*

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

so by the Copeland SCF, the winner is e. *However, notice that individual 1 has* e *at the bottom of his ranking. By being untruthful and completely reversing their reported ranking, individual 1 can sway the decision of the Copeland SCF to have* d *winning (with a score of 4) instead. This would be beneficial to individual 1 as they prefer* d *over* e

From the previous example we see that, in certain situations, an individual (in this case individual 1) is able to change the outcome of the Copeland SCF to their liking by reporting untruthful rankings. This means that the Copeland SCF is not "strategy proof". In the next section we will formalize this notion of being strategy proof and reveal a shocking result that states that strategy-proofness is in fact a relatively difficult property to have as a social choice function.

Gibbard-Satterthwaite Impossibility Theorem

The notion of strategy-proofness is an old one. In 1867, Charles Dodgson, better known as Lewis Caroll[4], was quoted saying

> "This principle of voting makes an election more of a game of skill than a real test of the wishes of the electors."

criticizing voting systems of the time. About a century later in 1973, both Allan Gibbard and Mark Satterthwaite each individually exploit Arrow's impossibility theorem to deduce a stronger result relating to strategy-proofness.

Definition 2.10. *An SCF, f is single voter manipulable if there exists two profiles P, P', and individual i such that $f(P') \succ_i f(P)$ and $\succsim_j = \succsim_j$ for all $j \neq i$; f is (single voter) strategy-proof if it is not single voter manipulable.*

In the case of the example with the Copeland SCF, we see that individual i was individual 1. They had the power to report untruthful rankings to change the profile from P to P' so that the outcome was more favourable to them. This means that Copeland SCF is single voter manipulable and not strategy-proof.

Definition 2.11. *An SCF f is nonimposed if, for every alternative in $x \in A$, there exists at least one profile P_x such that $f(P_x) = x$.*[5]

Definition 2.12. *An SCF f is resolute if, it has a single winner for all possible profiles. (there are no ties)*

Theorem 2.3. *An SCF is resolute, nonimposed, and strategy-proof for three or more alternatives if and only if it is dictatorial.*

Proof Outline:[6] We shall prove a simpler cas e assuming that there exists some SCF f is strategy-proof, resolute and pareto-optimal (stronger than nonimposing) and not a dictatorship. Consider the SCF f' that uses f to decide whether two alternatives a, b is preferred in the outcome. f' takes the profile P given to f and changes it to P' so that a and b is pushed to the top of every individuals ranking. ie. if an individuals ranking is $e \succ d \succ a \succ c \succ b$ it becomes $a \succ b \succ e \succ d \succ c$. f' then submits P' to f and checks the result to decide if $a \succ b$ or $b \succ a$. We can easily see that f' is an SCF that is not a dictatorship, pareto-optimal. With a little work one can prove that since f is strategy-proof and resolute, f' is also IIA. By Arrow's impossibility theorem, f' cannot exist, thus f must also not exist. □

[4]Yes, the one that wrote Alice in Wonderland. It turns out that he was also influential in the field of voting systems and elections. He went on to create the Dodgson Method, an extension to the Condorcet method to break Condorcet paradoxes

[5]Note that pareto-optimality implies nonimposition.

[6]adapted from Wikipedia page of Gibbard-Satterthwaite impossibility theorem

References

[1] A. Sen. *Collective Choice and Social Welfare: An Expanded Edition*. Harvard University Press, 2017.