Game Theory
Lecture #3

Focus of Lecture:

- Review social choice setup
- Arrow's Impossibility Theorem
- Proof

1 Social Choice

Last lecture introduced the problem of social choice and identified several desirable properties that any reasonable social choice mechanism should possess. As a reminder, the problem of social choice has the following key elements:

- Set of individuals: $N = \{1, \ldots, |N|\}$
- Set of alternatives: $X = \{x_1, \ldots, x_{|X|}\}$
- Preferences: For each individual $i \in N$ and every pair of alternatives $x, x' \in X$, exactly one of the following is satisfied:
  - $x \succ x'$ (i prefers x to x')
  - $x \prec x'$ (i prefers x' to x, which we will often write as $x' \succ x$)
  - $x \sim x'$ (i views x and x' as equivalent)

  We often will express the relations terms ($\succ$, $\prec$, $\sim$) as ($\succ_i$, $\prec_i$, $\sim_i$) to highlight the dependence on individual i. The preferences of the individuals contains a list of pairwise comparisons. For compactness, we sometimes will express the preferences of individual i by a function q_i where $q_i : X \times X \rightarrow \{\succ, \prec, \sim\}$ defines these pairwise preferences.

The goal in this social choice problem is to derive a social choice function $SC(\cdot)$ of the form:

$$q_N = SC(q_1, \ldots, q_n), \tag{1}$$

which takes in the preferences of the individuals and returns a single set of preferences of the form $q_N : X \times X \rightarrow \{\prec, \succ, \sim\}$. As we identified last lecture, several different social choice mechanisms yield results that are undesirable. Accordingly, we identified five fundamental properties (termed axioms) that any reasonable social choice mechanism should satisfy. As a reminder, these five axioms are as follows:

Axiom # 1 (Reasonable Domain and Range) *All preferences, both the individuals' preferences and societal preference, satisfy completeness and transitivity. Alternatively, all preferences can be expressed by a ranking.*

Axiom # 2 (Positive Association) *Improvements in the individual preferences of a given alternative should not degrade the societal preference of that particular alternative.*

Axiom # 3 (Unanimous Decision) *If all individuals prefer alternative x to y, then the societal preference should also prefer x to y.*

Axiom # 4 (Independence of Irrelevant Alternative) *The societal preference between a pair of alternatives is not impacted by the relative position of a third ("irrelevant") alternative.*

Axiom # 5 (Non-Dictatorship) *There should not be a dictator for any system with at least three individuals.*

The formal definition for each of these Axioms can be found in Lecture 2.

2 Arrow's Impossibility Result

With these axioms in place, our goal is to find a social choice mechanism that satisfies all five axioms. The following result from Kenneth Arrow in 1951 demonstrates that this goal is in fact mathematically impossible. That is, there are no social choice mechanisms that satisfy Axioms #1–5. The following theorem makes this statement precise.

Theorem 2.1 (Arrow, 1951) *If any social choice function $SC(\cdot)$ satisfies Axioms #1–4, then $SC(\cdot)$ necessarily does not satisfy Axiom #5.*

The remainder of this lecture is devoted to proving this impossibility result. In particular, we will start with a social choice function $SC(\cdot)$ that satisfies Axioms #1–4. We will not state the specific form of this social choice function. Rather, we will argue about the resulting social choice for various preference profiles $q = (q_1, \ldots, q_{|N|})$ using only Axioms #1–4. In doing so, we will identify several other properties that must be true for any such social choice function. Lastly, we will conclude the proof by showing that any social choice function $SC(\cdot)$ that satisfies Axioms #1–4 must have a dictator, in violation of Axiom #5.

The proof of Arrow's Impossibility Theorem will follow several steps. The starting point is a candidate social choice function, $SC(\cdot)$, that satisfies Axioms #1–4.

A central concept that we will use in the forthcoming proof is the idea of a *decisive set*, defined as follows:

Definition 2.1 (Decisive Set) *A set of individuals $V \subseteq N$ is decisive for the pair of alternatives (x, y) if for any preference profile $q = (q_1, \ldots, q_n)$ where $x \succ_i y$ for all $i \in V$, then the social choice $q_N = SC(q)$ also satisfies $x \succ_N y$.*

A decisive set identifies a collection of individuals V for each pair of alternatives (x, y) such that if all individuals $i \in V$ prefer $x \succ_i y$, then the societal choice $q_N = SC(q)$ also will satisfy $x \succ_N y$ irrespective of the preferences of the other individuals not in V. There are a two important observations regarding a decisive set:

(i) Each pair of alternatives (x, y) potentially could have a different decisive set V. Furthermore, the decisive set for (x, y) need not be the same as the decisive set for (y, x).

(ii) Since our social choice function $SC(\cdot)$ satisfies Axiom #3: Unanimous Decision, we know that N is a decisive set for each pair of alternatives (x, y).

The definition of a decisive set, along with Axiom #2: Positive Association and Axiom #4: Independence of Irrelevant Alternatives, together imply the following claim.

Claim 2.1 *Let $SC(\cdot)$ be a social choice function that satisfies Axioms #1, #2, and #4. Suppose there exists a preference profile q such that*

(i) $SC(q)$ has $x \succ_N y$ for some pair of alternatives (x, y);

(ii) individual j is the only agent with $x \succ_j y$; and

(iii) $x \prec_i y$ for all other individuals $i \neq j$.

Then the singleton set $\{j\}$ is decisive for the pair (x, y).

To prove this claim, let q' be any other preference profile with $x \succ_j y$ in q'. We will show that $SC(q')$ also results in $x \succ_N y$. If any agent $i \neq j$ has $x \succ_i y$ or $x \sim_i y$ in q', then by Axiom #2: Positive Association, $x \succ_N y$ in q', since at least one agent other than j improved the relative ranking of x versus y. Alternatively, if every agent $i \neq j$ kept $x \prec_i y$ in q', then by Axiom #4: Independence of Irrelevant Alternatives, it must be that $x \succ_N y$ in q' as well, since the relative ranking of x versus y for all agents is the same in q and q'. In both situations, $x \succ_N y$ in q'. Therefore, j is decisive for (x, y).

The following claim has a similar proof.

Claim 2.2 *Let $SC(\cdot)$ be a social choice function that satisfies Axioms #1, #2, and #4. Suppose there exists a preference profile q such that*

(i) $SC(q)$ has $x \succ_N y$ for some pair of alternatives (x, y);

(ii) for some $J \subset N$, $x \succ_j y$ for all $j \in J$; and

(iii) $x \prec_i y$ for all other individuals $i \notin J$.

Then the set J is decisive for the pair (x, y).

We now introduce the concept of a minimal decisive set:

Definition 2.2 (Minimal Decisive Set) *A set of individuals $V \subseteq N$ is a minimal decisive set if*

(i) there exists a pair of alternatives (x, y) such that V is decisive for the pair (x, y) and

(ii) any set $Q \subset V$ such that $|Q| < |V|$ is not decisive for any pair of alternatives (v, w).

The existence of at least one minimal decisive set is guaranteed. For any pair of alternatives (x, y), let us propose X as a candidate minimal decisive set. If the condition for minimality is violated by some subset $Q' \subset X$ and pair of alternatives, (v, w), then let Q be the new candidate minimal decisive set. By repeating this process, we can progressively whittle down to a minimal decisive set.

Note that any minimal decisive set, V, must satisfy $|V| > 0$, i.e., $V \neq \emptyset$. Observe that if $V = \emptyset$ was a decisive set for some pair of alternatives (x, y), this would imply that the social choice $q_N = SC(q_1, \ldots, q_n)$ would satisfy $x \succ_N y$ for any set of individual preferences $(q_1, \ldots, q_n)$. However, if $y \succ_i x$ for all individuals $i \in N$ and if the social choice mechanism $SC(\cdot)$ satisfies Axiom #3: Unanimous Decisions, then $y \succ_N x$ which is a contradiction. Hence, we have established the following key property:

Property #1 *Any social choice mechanism that satisfies Axioms #1–4 must have a minimal decisive set $V \neq \emptyset$ for some pair of alternatives (x, y).*

Now that we have established the existence of a minimal decisive set, V, we will seek to further identify structural properties about V. The remainder of the proof will repeatedly refer to the specific x from Property #1.

Since $V \neq \emptyset$, there exists at least one individual $j \in V$. Let $W = V \setminus \{j\}$ denote the remaining individuals in V. Accordingly, we have that $V = \{j\} \cup W$. We will probe the structure of the minimal decisive set V by analyzing the outcome associated with different preference profiles, q. To that end, consider a preference profile q of the form

$\{j\}$	W	U
x	z	y
y	x	z
z	y	x

where z is any alternative $\neq x, y$ and U is all individuals not in V. For clarity, all individuals in the set W have the ranking $z \succ x \succ y$, and all individuals in U have the ranking $y \succ z \succ x$.

What is the social choice associated with this preference profile, i.e., $q_N = SC(q)$? First note that the social choice q_N must satisfy $x \succ_N y$ as all individuals in $i \in V$ have $x \succ_i y$ and V is a decisive set for the pair (x, y). Second, suppose $z \succ_N y$. Observe that the only individuals $i \in N$ with preference $z \succ_i y$ is the set W with all others having $z \prec_i y$ for all

$i \notin W$. Hence, by Claim 2.2, W is decisive for the pair of alternatives (z, y). However, W cannot be a decisive set since W is a strict subset of V, and V is a minimal decisive set by assumption. Hence, we must have $y \succ_N z$ or $y \sim_N z$. Lastly, by transitivity we know that $x \succ_N z$.

Given that the resulting social choice $q_N = SC(q)$ must satisfy $x \succ_N z$, we observe that only one individual $\{j\}$ has a preference $x \succ_j z$. Hence, by Claim 2.1, individual $\{j\}$ is a decisive set for the pair (x, z). Since V is a minimal decisive set, it must be the case that $W = \emptyset$. This outcome leads to the following new property.

Property #2 *Any social choice mechanism that satisfies Axioms #1-4 must have a single individual* $\{j\}$ *that is decisive for any pair of alternatives* (x, z)*, with* $z \neq x$.

Note that the specific x here is the same one from Property #1.

Property #2 does not mean that individual $\{j\}$ is a dictator since that individual is only decisive for any pair of alternatives (x, z), with $z \neq x$. This means that if $x \succ_i z$, with $z \neq x$, then $x \succ_N z$ in the resulting social choice.

We will now enrich the set of examples we consider to show that individual $\{j\}$ is also decisive for other pairs of alternatives as well. To that end, let $z \neq x$ be any alternative and consider the following preference profile where U is all individuals not including $\{j\}$:

$\{j\}$	U
w	z
x	w
z	x

What can be inferred about the resulting social choice $q_N = SC(q)$ for this preference profile? First, $w \succ_N x$ since our social choice function $SC(\cdot)$ satisfies Axiom #3: Unanimous Decision. Furthermore, $x \succ_N z$ because individual $\{j\}$ is decisive for the pair of alternatives (x, z) (Property #2). Lastly, by transitivity we know that the resulting social choice must be of the form $w \succ_N x \succ_N z$. However, note that individual $\{j\}$ is the only individual that prefers $w \succ z$ and the resulting social choice satisfies $w \succ_N z$. Hence, by Claim 2.1, individual $\{j\}$ is also decisive for any pair of alternatives (w, z), with $w, z \neq x$. This conclusion leads to the following revised property.

Property #3 *Any social choice mechanism that satisfies Axioms #1-4 must have a single individual* $\{j\}$ *that is decisive for any pair of alternatives* (x, z)*,* $z \neq x$*, and* (w, v)*,* $w, v \neq x$.

Again, the x here is the same one from Property #1.

The last part of this proof entails showing that individual $\{j\}$ is also decisive for any pair of alternatives (z, x), $z \neq x$. To that end, let $w, z \neq x$ be any alternatives and consider the following preference profile where U is all individuals not including $\{j\}$:

$\{j\}$	U
w	z
z	x
x	w

What can be inferred about the resulting social choice $q_N = SC(q)$ for this preference profile? First, $z \succ_N x$ since our social choice function $SC(\cdot)$ satisfies Axiom #3: Unanimous Decision. Furthermore, $w \succ_N z$ because individual $\{j\}$ is decisive for the pair of alternatives (w, z), with $w, z \neq x$ (Property #3). Lastly, by transitivity we know that the resulting social choice must be of the form $w \succ_N z \succ_N x$. However, note that individual $\{j\}$ is the only individual that prefers $w \succ x$. Hence, by Claim 2.1, individual $\{j\}$ is also decisive for any pair of alternatives (w, x), with $w \neq x$. This last argument leads to the following final conclusion.

Property #4 *Any social choice mechanism that satisfies Axioms #1–4 must have a single individual $\{j\}$ that is decisive for any pair of alternatives (x, z), $z \neq x$, (w, v), $w, v \neq x$, and (z, x), $z \neq x$. Accordingly, individual $\{j\}$ is a dictator.*

This completes the proof.

3 Conclusion

This lecture focused on the design of social choice functions. Specifically, we asked the question of whether or not there are any *reasonable* mechanisms for aggregating the opinions of many? To that end, we identified five axioms that identified desired properties of any reasonable social choice mechanism. Our main result is the following amazing conclusion which is known as Arrow's Impossibility Theorem: *If any social choice function $SC(\cdot)$ satisfies Axioms 1-4, then the social choice function necessarily does not satisfy Axiom 5.* This is a famous result in economics that clearly illuminates the fundamental challenges associated with the design of social choice functions (or voting systems), as it is impossible to design a social choice function that meets our desired performance criteria. Accordingly, research has sought to identify what relaxations in our five axioms are necessary to ensure the existence of a social choice mechanism.

As an engineer, this negative result pertaining to social choice is an interesting cautionary tale. If merely aggregating societal opinions is hard (or provably impossible) then it could be the case that designing systems that will be utilized by society in an efficient manner is also hard or impossible. Given the emerging challenges faced by engineering in this interconnected world, it is extremely important for us to understand the limits of what is possible in these socio-technical systems.

4 Questions

1. Suppose a given social choice function satisfies Axioms #1-4. Further, suppose that the social choice function for a given preference profile is:

$$f\left(\begin{bmatrix} x & y & y \\ t & t & x \\ y & x & t \end{bmatrix}\right) = \begin{bmatrix} y \\ x \\ t \end{bmatrix}$$

   Can the social choice for the following preference profile be determined? If so, what is it?

$$f\left(\begin{bmatrix} x & x & y \\ t & t & x \\ y & y & t \end{bmatrix}\right) = ?$$

2. What are the decisive sets (with accompanying pairs) in the first preference profile in Question #1? What is the minimal decisive set? Assume the first column represents the preferences of individual 1, second column represents the preferences of individual 2, etc.

3. (a) Given the following preference profiles, can the social decision be determined when the guiding rule is to decide by a pairwise majority vote?

1	**2**	**3**
x	t	z
y	x	t
z	y	x
t	z	y

   (b) Now suppose that the social choice function satisfies Axioms #1-4. If it is also known that the social decision establishes a preference for y over z, is this information sufficient to predict the social decision?

   (c) Given the following preference profile and the information in (b), can the social decision be determined?

1	**2**	**3**
x	z	z
y	x	t
z	y	x
t	t	y