Math 7h Professor: Padraic Bartlett

Lecture 3: Democracy and Arrow's Theorem

Week 3 *UCSB 2014*

Consider the following problem, encountered by democracies all over the world:

Problem. Suppose that you have a set of **choices** $\mathbb{A} = \{\alpha, \beta, \gamma, \delta \ldots\}$. Given such a set, a **ranking** of $\mathbb{A}$ is simply some ordering of the elements of $\mathbb{A}$: for example, one ranking of the set {Yeats, Emerson, Blake} could be

$$\text{Yeats} \ > \ \text{Blake} \ > \ \text{Emerson}.$$

Call the collection of all possible rankings $\mathcal{R}$.

A **voting system** C on N voters is simply any function $C : \mathcal{R}^N \to \mathcal{R}$. In other words, it is a function that takes in any set of N rankings, and uses these rankings to determine some overall "social preference."

Some voting systems are better than others. For example, under the choice set { cake, pie, custard, all-encompassing doom}, a plausible voting system could be the map $C : \mathcal{R}^N \to \mathcal{R}$,

$$C(R_1, \ldots R_N) = (\text{all-encompassing doom} \ > \ \text{cake} \ > \ \text{pie} \ > \ \text{custard}).$$

In other words, our voting system takes in all of the preferences of our voters, completely ignores those preferences, and instead selects doom as its top-ranked choice (with cake, pie and custard ordered after doom.)

This is . . . not ideal. In theory, it would be nice if our voting system reflected, in some way, the desires of our voters! This raises the following question: what are the properties we want in a voting system?

After some thinking, a legislative body might come up with the following desired preferences:

1. **Unanimity**: Suppose that every member of our society submits ballots where choice α is ranked above choice β. Then our voting system C should reflect this choice as well.

   In other words, if $(R_1, \ldots R_N) = \vec{R} \in \mathcal{R}^N$ is a collection of rankings such that $\alpha > \beta$ for each R_i, then $\alpha > \beta$ should hold for $C(\vec{R})$ as well.

2. **Independence of Irrelevant Alternatives**: Suppose we have two options α and β, and the only thing we are concerned with is whether α is ranked above or below β in the output of our voting system. Then the only information that "matters" for deciding this information should be each voter's relative ranking of α to β.

   In other words: suppose you hold one election that results in the outcome $\alpha > \beta$. If we hold a second election where each voter's preference between α and β doesn't change — i.e. if you used to support α over β, you still do, and vice-versa — but maybe your preferences for some other options moves around (i.e. δ moves to the top of your list.) This shouldn't change our results: after all, if no-one's preferences between α and β changed, why should changing irrelevant information matter?

Formally speaking, this is the following claim: suppose we have two vectors $\vec{R}, \vec{S}$ such that for each i, the rankings R_i, S_i both agree on the relative ranking of α to β. Then the relative ranking of α to β in the two results $C(\vec{R})$, $C(\vec{S})$ should be the same.

We call any voting system that satisfies these two properties above "fair." What are some "fair" voting systems? Well: if we only have two options we can simply use **majority rule**:

Voting System.
(Majority Rule.) Given any collection $\vec{R} \in \mathcal{R}^N$ of rankings on a two-choice set $\mathbb{A} = \{\alpha, \beta\}$, we can define the voting system $C(\vec{R}) \to \mathcal{R}$ as follows:

- If more rankings have $\alpha > \beta$ than the other way around, output $\alpha > \beta$.
- Otherwise, output $\beta > \alpha$.

This test passes the **unanimity** condition (because if everyone prefers α to β, those ballots will outnumber the $\beta > \alpha$ ballots trivially) and the **independence of irrelevant alternatives** condition (because there are no other alternatives to consider.)

For multiple-choice systems, though, the idea as above won't work literally as written, as it doesn't tell us what to do about our non-α, β choices! So: what is a simply-defined multiple-choice voting system? Well, one approach that has been (sadly) popular throughout history is the following:

Voting System.
(Dictatorship.) Take any collection $\vec{R} \in \mathcal{R}^N$ of rankings on a choice set $\mathbb{A}$. As well, call one voter i the **dictator**. Then, we can define the voting system function $C(\vec{R}) \to \mathcal{R}$ as follows:

$$C(\vec{R}) = R_i.$$

In other words, this just looks up what voter i's preference is and outputs that preference!

Surprisingly, this system satisfies the two requirements of **unanimity** and **irrelevance of independent alternatives** that we asked our systems to satisfy above. If every voter ranks $\alpha > \beta$, then in particular our dictator preferred α to β, and therefore in our output we have $\alpha > \beta$: i.e. we satisfy unanimity. Similarly, if don't change the relative ranking of α to β in anyone's vote, then in particular we don't change the relative ranking of α to β in our dictator's vote, and thus we satisfy the irrelevance of independent alternatives condition.

Hmm. Can we do better? Perhaps more generally: what kinds of properties do voting systems have to have? Can we use mathematics to deduce certain properties about how these things work?

As you may have guessed, given that we're talking about this in a mathematics class: yes! We start with the following lemma:

Lemma. (The extremal lemma.) Suppose we have some collection of choices $\mathbb{A}$, where $\mathbb{A}$ contains at least three different choices. Pick any choice $\alpha \in \mathbb{A}$, and suppose we have a collection of votes $\vec{R}$ such that in each vote R_i, either α is at the top of the ranking R_i or at the bottom of the ranking R_i.

Suppose that $C : \mathcal{R}^N \to \mathcal{R}$ is a "fair" voting system. Then in the ranking $C(\vec{R})$, the choice α must either be greater than every other choice, or smaller than every other choices.

Proof. We will proceed by contradiction: in other words, we will suppose for the moment that there was a collection of votes $\vec{R} = (R_1, \dots R_n)$ in which α was always at the top or the bottom of each vote R_i, and yet somehow in $C(\vec{R})$ we have $\beta > \alpha > \gamma$ for two other options β, γ.

Consider each vote R_i. Notice that by definition we know that α is always at the top or the bottom of each person's vote. Now, suppose that we take each vote R_i, and modified it by placing γ directly above β and moving everything else down one. In other words, if we had the vote

$$(\alpha > \delta > \sigma > \boxed{\beta} > \theta > \boxed{\gamma} > \phi),$$

we would replace it with the vote

$$(\alpha > \delta > \sigma > \boxed{\gamma} > \boxed{\beta} > \theta > \phi).$$

Call this modified collection of votes $\vec{R}'$. Notice that because α is always at the exact top of bottom of our list, the relative position of α to every other option never changes. Therefore, by the independence of irrelevant alternatives, because we never changed the relative ranking of α to β or the relative ranking of α to γ, their rankings in the output of C never changed: i.e. we **still** get $\beta > \alpha > \gamma$ in $C(\vec{R}')$.

But in $\vec{R}'$, we have $\gamma > \beta$ on **everyone**'s vote! Therefore, by the unanimity condition we must have $\gamma > \beta$ in our result $C(\vec{R}')$, which contradicts our claim that $\beta > \alpha > \gamma$.

Consequently we have found a contradiction! In other words, if our voting system is fair, and each individual vote ranks α either first or last, then our voting system must also rank α either first or last. ☐

This lemma has an interesting corollary:

Corollary. Take any two options α, β. Then is a collection of votes $\vec{R}$ and a "pivotal" voter i such that the vote R_i of voter i determines entirely what happens to α: i.e. by changing R_i, we can put α at either the top or the bottom of the ranking $C(\vec{R})$.

Proof. To see this, let $\vec{R}$ denote any collection of votes such that in each ranking R_i, α is at the top of this ranking. By unanimity, $C(\vec{R})$ must also have α at the top of its ranking. Now, one-by-one take each of the rankings $R_1, R_2, R_3 \dots$ and move α to the bottom of each ranking. At the end of this process, every voter has α at the bottom of their rankings, and consequently $C(\vec{R})$ must have α at the bottom of its rankings. Note that along the way α is always at either the top or the bottom of $C(\vec{R})$: this is because of our lemma, as α is at the top or the bottom of each individual voter's ranking.

Therefore, at some point in time, there was a voter k_α such that when k_α changed their vote R_{k_α}, $C(\vec{R})$'s rankings changed from having α at the top to α at the bottom. This is exactly the voter we were looking for! ☐

For future reference, call any collection of votes where $R_1, R_2 \dots R_{k_\alpha - 1}$ all have α at the bottom of their rankings and $R_{k_\alpha+1}, \dots R_n$ have α at the top of their rankings (k_α, α)-**critical**. In this sense, we have proven that in any (k_α, α)-**critical** collection of votes, the voter k_α is pivotal: their vote determines whether α is at the top or the bottom in $C(\vec{R})$!

Let's consider this "pivotal" voter k_α. Do they have a lot of influence over other events in our election? In other words: can they use this "influence" over α to change the rankings of other events?

Well: take any (k_α, α)-critical collection of votes $\vec{R}$, and consider any two other options $\beta, \gamma \neq \alpha$. Suppose that we let all of our voters move β and γ around in their orderings, while preserving the "top or bottom" placement of α. By the independence of irrelevant alternatives, this doesn't change α's position at the top or bottom of our rankings, and furthermore does not change the fact that k_α's vote determines whether α is at the top or the bottom in $C(\vec{R}_\alpha)$. In particular, if R_{k_α} has α as its top-ranked choice, we would have $\alpha > \beta$ and $\alpha > \gamma$ in $C(\vec{R})$; similarly, if R_{k_α} has α as its bottom-ranked choice, we would have $\beta > \alpha$ and $\beta > \gamma$ in $C(\vec{R})$.

Now, suppose that voter k_α rearranges their rankings in some fashion and puts $\beta > \alpha > \gamma$. By the independence of irrelevant alternatives, we must have $\alpha > \gamma$. This is because if we compare our collection of votes to the collection where α was at the top of k_α's vote, we have not changed the relative position of α to γ.

Similarly, by the independence of irrelevant alternatives, we must have $\beta > \alpha$. This is because if we compare our collection of votes to the collection where α was at the bottom of k_α's vote, we have not changed the relative position of β to α.

Therefore, in a sense, the only thing that mattered was k_α's vote! I.e. by placing α, the voter k_α can declare that everything above where they place α will finish ahead of everything below where they place α. In other words, we have the following result:

Lemma. Take any (k_α, α)-critical collection of votes $\vec{R}$, and let $\beta, \gamma \neq \alpha$ be a pair of non-α options. Then $\beta > \alpha > \gamma$ in R_{k_α} if and only if we have $\beta > \alpha > \gamma$ in $C(\vec{R})$ as well.

In particular, this gives us the following corollary:

Corollary. Take any collection of votes $\vec{R}$, and let $\beta, \gamma \neq \alpha$ be a pair of non-α options. Then $\beta > \gamma$ in R_{k_α} if and only if $\beta > \gamma$ in $C(\vec{R})$ as well.

Proof. Take our collection of votes $\vec{R}$. Move α to the highest slot in each vote $R_1, \dots R_{k_\alpha - 1}$, and to the lowest slot in each vote $R_{k_\alpha+1}, \dots R_n$. Then the lemma above tells us that $\beta > \alpha > \gamma$ in R_{k_α} if and only if we have $\beta > \alpha > \gamma$ in $C(\vec{R})$.

But if we move α back to where it originally was in $\vec{R}$, this doesn't change the relative positioning of β to γ! Therefore, we actually have that $\beta > \gamma$ in R_{k_α} if and only if $\beta > \gamma$ in $C(\vec{R})$ as well. $\square$

This seems … bad. How bad?

Theorem. The only fair voting system in a situation where there are three or more options is a dictatorship.

Proof. take some option $\delta \neq \alpha$, and use our first lemma to find a (k_δ, δ)-critical collection of votes $\vec{R}$. Now, take $\vec{R}$, and move α to the highest non-δ slot possible for each vote $R_1, \ldots R_{k_\alpha - 1}$, and to the lowest non-$\beta$ slot possible for each vote $R_{k_\alpha+1}, \ldots R_n$.

Therefore, if you "ignore" δ, this is a (kind-of) (k_α, α)-critical collection as well! In particular, by the irrelevance of independent alternatives, we know that the corollary we just mentioned above still goes through: i.e. if we pick any $\beta, \gamma \neq \alpha, \delta$, we will have $\beta > \gamma$ in R_{k_α} if and only if $\beta > \gamma$ in $C(\vec{R})$.

But this corollary **also** tells us that $\beta > \gamma$ in R_{k_δ} if and only if $\beta > \gamma$ in $C(\vec{R})$. So, in particular, if $k_\delta \neq k_\alpha$, we could create a contradiction by placing $\beta > \gamma$ in R_{k_α} and $\gamma > \beta$ in R_{k_δ}! This cannot happen: therefore, we must have $k_\alpha = k_\delta$!

Call this unique pivot k. We now have that for **any** two options β, γ we have $\beta > \gamma$ in R_k if and only if $\beta > \gamma$ in $C(\vec{R})$, by simply thinking of $k = k_\alpha$, for some $\alpha \neq \beta, \gamma$. In other words, k is a dictator, and our voting system is a dictatorship. $\square$