Democratic Elections and Arrow's Impossibility Theorem

Brendan Lauterborn

12/7/2019

Arrow's Impossibility Theorem identifies that, for any election consisting of three or more candidates, there is no ideal voting system that meets all the reasonable properties. Without first understanding the axioms set by Arrow, it is impossible to recognize why there exists no democratic election that can satisfy its intended purpose. Jean Charles de Bord, Charles Dodgson and Duncan Black have also studied this problem with less success.

Arrow introduces a preferred voting method, known as the social welfare function, by which voters rank all of the candidates in order of preference. After collecting the results, the system produces an output that represents the group decision. Unfortunately, no system produces a group decision without breaking one of the five properties, which are Unrestricted Domain, Pareto Efficiency, Independence of Irrelevant Alternatives, Non-Dictatorship, and Social Ordering.

The first property is Unrestricted Domain. The definition states that, "The domain of f includes every list $\langle R_1, \ldots, R_n \rangle$ of n weak orderings of X" (Morreau). This definition means that any ordering of preferences is allowed. The voter is given the ability to rank the candidates in any order that they wish to. Let X, Y and Z represent three voters and A, B and C represent three candidates. Define $X : A \succ B \succ C$, $Y : B \succ C \succ A$ and $Z : C \succ B \succ A$. This statement says that voter X prefers candidate A to B and B to C, and so on. This ordering is allowed. There is no restriction on the ordering of candidates.

The second axiom is Independence of Irrelevant Alternatives. The formal definition is, "For all alternatives x and y in X, and all profiles $\langle R_i \rangle$ and $\langle R_i^* \rangle$ in the domain of f, if $\langle R_i \rangle|_{\{x,y\}} = \langle R_i^* \rangle|_{\{x,y\}}$, then $f(\langle R_i \rangle)|_{\{x,y\}} = f(\langle R_i^* \rangle)|_{\{x,y\}}$" (Morreau), where alternatives represent a different possible candidate that one can vote for and the profile represents how each voter ranks each of the candidates. The statement $\langle R_i \rangle|_{\{x,y\}} = \langle R_i^* \rangle|_{\{x,y\}} \Rightarrow f(\langle R_i \rangle)|_{\{x,y\}} = f(\langle R_i^* \rangle)|_{\{x,y\}}$ means that, whenever two profiles $\langle R_i \rangle, \langle R_i^* \rangle$ are identical, the social preference relations $f(\langle R_i \rangle), f(\langle R_i^* \rangle)$ must be as well. In other words, this definition states that if society as a whole prefers one candidate to another, introducing a new candidate is irrelevant to the outcome. Similarly, if one candidate is preferred over another by all people, taking away the least preferred candidate is irrelevant. For example, let A and B be candidates and $A \succ B$ by all people. To satisfy Independence of Irrelevant Alternatives, introducing or removing C should not affect the preference between A and B. This definition holds whether C is placed at the very top or the very bottom.

The third axiom is Non-Dictatorship. This declares that the results of an election cannot agree with a single individual. The definition states that a social welfare function is dictatorial if, "there exists an individual i such that, for all x and y, $x \succ_i y$ implies $x \succ y$ regardless of the orderings $R_1, \ldots, R_n$ of all individuals other than i" (Arrow), where $x \succ_i y$ means that candidate x is preferred to candidate y for some voter i, and similarly, $x \succ y$ means that candidate x is preferred to candidate y for all people. To reiterate, a single voter's decision cannot overrule the majority decision of all other voters. Let voters X, Y prefer $A \succ B$ while voter Z prefers $B \succ A$. The resulting conclusion must be that $A \succ B$. Voter Z cannot overrule the two other voters.

The next property is known as Social Ordering. The definition states, "For any profile $\langle R_i \rangle$ in the domain, $f(\langle R_i \rangle)$ is a weak ordering of X" (Morreau), where f is the welfare function. f is used in the formal definition, but I will be using $\succ_W$ to refer to the social welfare function once I begin the proof. The definition means that, "the result of aggregating individual preferences is always a weak ordering of the alternatives" (Morreau). This definition implies that the ordering of preferences has to be from better to worse. There can possibly be ties, ultimately ruling out the paradox of voting, which states that there can be a cycle of preferences. To show the paradox of voting or Condorcet's paradox, let X, Y and Z be voters and A, B and C be candidates. Suppose $X : A \succ B \succ C$, $Y : B \succ C \succ A$ and $Z : C \succ A \succ B$. From this statement, we are able to see that candidate A is preferred to B and candidate B is preferred to C. However, candidate C is also preferred to A. Social Ordering rules prevent a cycle from occurring.

The final axiom is Pareto Efficiency. The definition states that, "For any profile $\langle R_i \rangle$ in the domain of f, and any alternatives x and y, if for all i, $x \succ_i y$, then $x \succ y$" (Morreau). This definition means that for all voters i, if candidate A is preferred to candidate B, then the social welfare function must also prefer candidate A to candidate B. In other words, if every voter agrees on the ordering of two candidates, then society does as well. Let $1, 2, 3, \ldots, n$ be individual voters and A, B be candidates. Suppose 1 prefers $A \succ B$, 2 prefers $A \succ B$, 3 prefers $A \succ B$, $\ldots, n$ prefers $A \succ B$. Then it is true that society prefers $A \succ B$. Another thing we assume is transitivity. Transitivity states that, "For all x, y and z, xRy and yRz imply xRz" (Arrow), where xRy reads, x is preferred or indifferent to y. Now that the conditions are set, we can begin looking at why there is no social welfare function that satisfies all of the conditions.

Much of the following proof was drawn out by Dr. Kevin Leyton-Brown and Dr. Eric Pacuit. I will simply reiterate what they had to say and elaborate on the interesting points. Arrow's Impossibility Theorem states that there is no social welfare function that satisfies all five axioms. For simplicity, I will mainly focus on a welfare function that is Pareto Efficient and Independent of Irrelevant Alternatives. So, what I am trying to prove is that, "Any social welfare function that is Pareto efficient and independent of irrelevant alternatives is dictatorial" (Arrow).

We seek to prove that this is true. We will do this by showing that there exists a pivotal voter who can alter the outcome of the social welfare function, thus making them a dictator.

There will be four parts to the proof. The first claim is that if every voter places a candidate b at either the very top or the very bottom, then our social welfare function must do the same. The second part of the proof is to show that there exists some voter, n', such that by altering his/her vote, they will change the outcome b from the bottom of the welfare function to the top. The third claim is that the voter n' is a dictator over any pair a, c that does not involve b. The final part is to show that this voter n' is a dictator over all pairs a, b.

With that being said, we will assume that our social welfare function is both Pareto Efficient and Independent of Irrelevant Alternatives. We also assume that there are three or more outcomes, or candidates, in our function. Suppose there is a ranking, R, that contains some number of voters who place outcome b at the very top or the very bottom; placing a and c wherever. Our first claim is that if every voter places an outcome b at either the very top or the very bottom, then our social welfare function must do the same.

Let us assume that our claim is false. That is, we assume that social welfare function places b in the middle between a and c. Suppose we have another ranking, R', that contains some number of voters who place b at either the very top or the very bottom. This assumption implies that a is either above or below b, depending on the voter. Now, if I place c directly above a, leaving b unchanged, we can have the following: $c \succ a \succ b$ or $b \succ c \succ a$. By definition of independence of irrelevant alternatives, we know that in order for $a \succ_W b$ or $b \succ_W c$ to change, the relationship between a, c or the relationship between b, c would need to change. However, since b occupies an extremal position for all voters, c can be moved above a without changing either of these pairwise relationships (Leyton-Brown). We see that in R' the social welfare function says that $a \succ_W b$ and $b \succ_W c$. Recall that transitivity declares that if $a \succ b$ and $b \succ c$, then $a \succ c$. However, we see that in the ranking R', every single voter ranks c above a. By Pareto Efficiency, the social welfare function must also rank c above a. However, our social welfare function of R' states $a \succ_W c$. Thus, we have a contradiction. We see that the claim has to be true.

This part of the proof is straightforward. We can see that if everyone in the room says that they prefer dogs to cats and birds, then it must be true that the social welfare function prefers dogs to cats and birds. Similarly, if everyone in the room says that their least favorite pet is a dog, then the welfare function must also reflect that. By seeing and understanding this point, it helps us to set up the next part of the proof. Our second claim is that there exists some voter n' such that by altering his/her vote, they can change the outcome b from the bottom of the welfare function to the top.

Suppose there is a ranking profile R_0 where $1, \ldots, n$ voters rank b to be last. The profile R_1 has voter 1 move b to the very top of the list. Voters $2, \ldots, n$ remain unchanged. Profile R_2 now has voter 1 and voter 2 move b to the very top of the list. Voters $3, \ldots, n$ remain unchanged. The social profile R_m has voters $1, \ldots, n$ rank b at the very top. For each profile R, b is either at the very top or the very bottom. Notice that for R_0, by Pareto Efficiency, b cannot be ranked at the very top. Similarly, for R_m, again by Pareto Efficiency, b cannot be ranked at the very bottom. This means that as we move along our profiles, there must be a time where the position of b is altered from the bottom to the top in the welfare function. We denote n' to be this voter that alters the welfare function from having b at the bottom, to having b at the top. Therefore, we are able to see that n' can alter the social preference by changing his/her vote.

This part of the proof is interesting. We go from a somewhat simple step to a fairly large claim. We are able to prove that there exists a single person that is able to alter the placement of b in an election. If this voter, n', places the candidate at the very bottom, the welfare function does so as well. If this voter places the candidate at the very top, the welfare function does so as well here also. Thus, this voter has a lot of power. This part of the proof shows that the other voter's preferences of b don't mean anything. The dictator essentially throws them away. The next part of the proof is even more interesting because not only is n' a pivotal voter for b, but they are also a dictator over any pair that does not involve b. Our claim is that voter n' is a dictator over any pair a, c that does not involve b.

Suppose we have some ranking profile R_0 that has every voter up to $n'-1$ rank b at the very top. Voter n' to n has b ranked at the very bottom. Each voter places a and c wherever they choose. This means that voter 1 can have $a \succ c$, $c \succ a$ or be indifferent between the two. We have another ranking profile R_1 that has every voter up to n', inclusive of n', rank b at the very top. Voters $n'+1$ to n rank b at the very bottom. Again, each voter places a and c wherever they choose. Now, we create a third profile R_2. R_2 is constructed the exact same way as R_1, but we alter n''s preferences. We move a to the very top and leave b and c where they were. The preference of n' is now $a \succ b \succ c$. Notice that in R_0 and R_2, $a \succ b$ because n' has now ranked a above b and the following voters too have candidate a over b. By Independence of Irrelevant Alternatives, the social welfare function in R_2 also prefers a to b. Similarly, if we compare R_1 and R_2, we see that $b \succ c$ in both. By definition of Independence of Irrelevant Alternatives, we know that $b \succ_W c$ in the profile R_2. By definition of transitivity, it must be true that $a \succ_W c$ in R_2. Let us now create a fourth ranking profile R_3. In R_3, each voter other than n' keeps the same ordering that they had before, except they move b to some different position without altering the other preferences. Voter n' now moves b below a, but higher than c. In R_2 and R_3, the voters have the same preferences between a and c. Therefore, since we know that $a \succ_W c$ in R_2, and by Independence of Irrelevant Alternatives, then $a \succ_W c$ in R_3. Thus, "we have determined the social preference between a and c without assuming anything except $a \succ_{n'} c$" (Leyton-Brown). This proves that the claim is true. Voter n' is a dictator over a, c.

From this step, we are able to show that this single voter can determine how the pair a, c is placed, simply by assuming that voter n' prefers a to c. One single voter determines how the social welfare function places a and c. This illustrates that the voting system is indeed flawed. There should never be a system where the outcome is determined by a single person; however, we see that one exists. The final part of the proof is to prove that n' is a dictator over any pair. Our claim is voter n' is a dictator over all pairs, a, b.

Let c be the third outcome. In the second part of the proof, we saw that n' was able to alter the social welfare function by altering their preference of b. We see that there is a voter n'' that can do the same for c, where n'' is the pivotal voter. By the results of the third step in the proof, we know that n'' is a dictator over alternatives other than c. We now seek to show that $n'' = n'$. Suppose n'' is placed ahead above n' in the profile R_0. That is, n' has b still at the very bottom. Voter n'' has b at the very top and a right below it. It must be true that if n'' ranks $b \succ a$ then the welfare function must follow. However, that does not happen. In this profile, $a \succ_W b$. n' is still the pivotal voter. Let's consider that n'' is placed after n' in R_1. In this profile, the welfare function has $b \succ_W a$. Being that n'' is placed after n', we have $a \succ_{n''} b$. Therefore, n'' cannot be the dictator in this case. Thus, voter $n' = n''$. This means that n' is a dictator over all alternatives.

This is remarkable. Through multiple claims, it was proven that there can only be a single dictator and that this dictator determines the ordering of all pairs. The outcome of an election depends only on the preferences of this dictator, not other people. This contradicts the intent of a democratic voting system.

We conclude that Arrow's Impossibility Theorem is true. I find it very interesting that no proper democratic voting system exists given a set of just rules. Assuming unrestricted domain, pareto efficiency, independence of irrelevant alternatives and social ordering, the social welfare function must be dictatorial.

References

[1] Arrow, Kenneth J. *Social Choice and Individual Values.* 1951.

[2] "Arrow's Impossibility Theorem." *Math Fun Facts.* `https://www.math.hmc.edu/funfacts/ffiles/20005.8.shtml`

[3] "Arrow's Impossibility Theorem." *Why Do Math?* `http://www.whydomath.org/node/voting/Arrow's_Impossibility_Theorem.html`

[4] Fey, Mark. *A Straightforward Proof of Arrow's Theorem.* 2014.

[5] Hayden, Grant M. "Some Implications of Arrow's Theorem for Voting Rights." *Stanford Law Review*, 1995. https://www.jstor.org/stable/1229229

[6] Leyton-Brown, Kevin. *Social Choice: Arrow's Theorem.* 2013. https://www.youtube.com/watch?v=QLi_5LCwJ20

[7] McComb, Christopher et al. *Design Science: Impossible by Design? Fairness, Strategy, and Arrow's Impossibility Theorem.* Cambridge University Press, 2017.

[8] Morreau, Michael. "Arrow's Theorem." *Stanford Encyclopedia of Philosophy.* 2014. https://plato.stanford.edu/entries/arrows-theorem

[9] Pacuit, Eric. *Proof of Arrow's Theorem.* 2015. https://www.youtube.com/watch?v=LVkioI5Z0JE