V. Dardanoni's Pedagogical proof of Arrow's Impossibility Theorem

Robin Whitty, Maths Study Group,
December 19th 2024

This talk is largely based on
Valentino Dardanoni, "A pedagogical proof of Arrow's Impossibility Theorem",
Social Choice and Welfare, Vol. 18, No. 1, 2001, pp. 107–112.

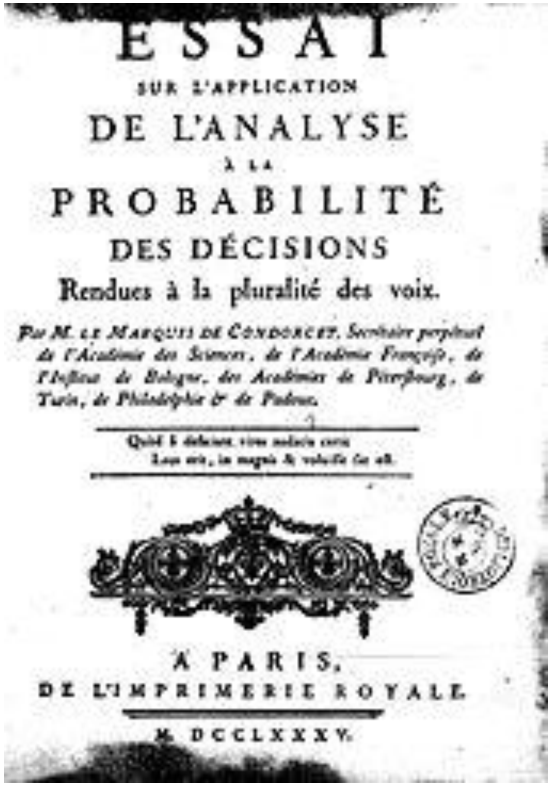
Condorcet’s Paradox 1

Suppose you have an election with 3 candidates X, Y and Z.
Suppose that, by a majority, X outvotes Y and Y outvotes Z.
You would expect that X should outvote Z.

$$X > Y \text{ and } Y > Z \text{ presupposes that } X > Z$$

En 1785, Nicolas de Condorcet discovered that, paradoxically, such a comparison of candidates by majority voting could be **non-transitive**.

Marie Jean Antoine Nicolas de Caritat,
marquis de Condorcet (1743 – 1794)



ESSAI
SUR L'APPLICATION
DE L'ANALYSE
À LA
PROBABILITÉ
DES DÉCISIONS
Rendues à la pluralité des voix.

Par M. le Marquis de Condorcet, Secrétaire perpétuel de l'Académie des Sciences, de l'Académie Françoise, de l'Institut de Bologne, des Académies de Pétersbourg, de Turin, de Philadelphie & de Padoue.

A PARIS,
DE L'IMPRIMERIE ROYALE.
M. DCCLXXXV.

Condorcet's Paradox 2

In the table below, X, Y et Z, are compared by three voters, A,B et C.

The table shows all the configurations of comparisons between each pair of candidates.

For now we just look at one configuration from each part of the table.

	X vs Y									Y vs Z									Z vs X							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**		**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**		**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
A	>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<

The majority votes are NON-TRANSITIVE ! $X > Y$ and $Y > Z$ but $Z > X$

Nevertheless the three voters behave **rationally**, i.e. transitively.

E.g. A : X < Y, which is consistent with Y > Z, Z > X, etc.

Note that lots of combinations of configurations are not rational, e.g. columns 1,9 and 17.

Arrow’s Impossibility Theorem 1

In 1948 Kenneth Arrow constructed a profound extension of Condorcet’s paradox. Condorcet shows that an intuitively desirable property of voters, rationality, need not be preserved under majority voting.

Kenneth Joseph Arrow (1921 – 2017)

In the same spirit, Arrow proposed some conditions that seemed desirable for a democratic election, and demonstrated that they inevitably lead to a dictatorship.

A ‘dictator’ means a voter whose choices between candidates are always mirrored by the ‘algorithm’ that amalgamates the votes for the candidates.

A dictator is not necessarily someone who exploits the electoral process to seize and hold on to power. They many not even realise that their choices mirror the algorithm used by the electoral process. (But if they do they could presumably act undemocratically).

Arrow’s ideas, set out in the 1951 book “Social Choice and Individual Values”, gave birth to a whole branch of economic, Social Choice Theory, and earned him the Nobel Prize in 1972.

A PEDAGOGICAL PROOF OF ARROW'S IMPOSSIBILITY THEOREM

VALENTINO DARDANONI, UNIVERSITÀ DI PALERMO

ABSTRACT

In this note I consider a simple proof of Arrow's Impossibility Theorem (Arrow 1963). I start with the case of three individuals who have preferences on three alternatives. In this special case there are $13^3 = 2197$ possible combinations of the three individuals' rational preferences. However, by considering the subset of *linear* preferences, and employing the full strength of the IIA axiom, I reduce the number of cases necessary to completely describe the SWF to a small number, allowing an elementary proof suitable for most undergraduate students.

	X vs Y									Y vs Z									Z vs X							
	1	2	3	4	5	6	7	8		9	10	11	12	13	14	15	16		17	18	19	20	21	22	23	24
A	>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<

(Here there are 8^3 combinations of three columns, one from part of the table (comparison of candidates) , although columns 1, 9 and 17, say, are not rational (transitive). Dardanoni allows, in principle, five comparisons: $\leq, \geq, <, >, =$, which seems to give 125^3 combinations, so I'm not sure where 13^3 comes from.)

A PEDAGOGICAL PROOF OF ARROW'S IMPOSSIBILITY THEOREM

VALENTINO DARDANONI, UNIVERSITÀ DI PALERMO

Suppose we have 3 individuals, a, b and c, who have preferences on 3 alternatives, x, y and z. Denote weak preference by $\succeq_i$, strict preference by $\succ_i$ and indifference by $\sim_i$, $i = a, b, c, s$, with s denoting society. A rational preference relation is complete and transitive (a weak ordering). A preference relation is *linear* when it is also antisymmetric, so that no two distinct alternatives are ever indifferent. Note that preferences can always be described by considering the ranking in the three *pairwise comparisons* x versus y, y versus z and z versus x.

	X vs Y									Y vs Z									Z vs X							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**		**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**		**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
A	>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<

This special case conveys the nature of Arrow's result. It is well known that the restriction to three options is not really limiting (any larger set of alternatives can be broken down into triplets, and any inconsistency within a triplet implies an inconsistency on the larger set). However, the general case of $n \geq 3$ individuals can be easily considered in this framework, by building on the proof of the simpler case. I hope that a motivated student, having mastered the simple case of three individuals, will find this extension approachable and rewarding.

Arrow's Impossibility Theorem 2

Going back to Condorcet :

	X vs Y									Y vs Z									Z vs X							
	1	2	3	4	5	6	7	8		9	10	11	12	13	14	15	16		17	18	19	20	21	22	23	24
A	>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<

Majorities : > > > non-transitive !

Arrow required that voters should be **rational**, meaning their choices should be transitive.
In the configurations below, A is being irrational.

	X vs Y									Y vs Z									Z vs X							
	1	2	3	4	5	6	7	8		9	10	11	12	13	14	15	16		17	18	19	20	21	22	23	24
A	>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<

(Condorcet showed, nevertheless, that the choices of rational voters could be combined to give an irrational election outcome.)

A PEDAGOGICAL PROOF OF ARROW'S IMPOSSIBILITY THEOREM

VALENTINO DARDANONI, UNIVERSITÀ DI PALERMO

Definitions: *A Rational Unrestricted-Domain Social Welfare Function (SWF) is a function that takes any three rational individual preferences and gives back a rational social preference. A SWF satisfies Independence of Irrelevant Alternatives (IIA) when social ranking on a given pairwise comparison depends only on individuals' ranking on that comparison. A SWF satisfies Unanimity (U) if, whenever everybody has the same strict ranking on a given pairwise comparison, then society has the same ranking. Individual i is a dictator if society's preference always coincide with individual's i strict preference regardless of all the other individuals' preferences. A SWF satisfies Non-Dictatorship (ND) if there is no dictator.*

X vs Y

	1	2	3	4	5	6	7	8
A	>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<
SWF:					**>**			

Y vs Z

	9	10	11	12	13	14	15	16
A	>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<
SWF:			**>**					

Z vs X

	17	18	19	20	21	22	23	24
A	>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<
SWF:		**<**						

So the Social Welfare Function is a function like

$$SWF:\{1,\ldots,8\}\times\{9,\ldots,16\}\times\{17,\ldots,24\}\rightarrow\{<,>\}^3$$

except not all triples in the domain represent rational voters and not all triples in the codomain are transitive.

Arrow’s Impossibility Theorem 3

Next, Arrow modelled an election with a table like the one we’re using, with each **collection of 3 columns**, one from each part of the table, being a potential election, called a **preference profile.**

	X vs Y								Y vs Z								Z vs X							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
A	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<

1. **Arrow’s model**: an election is a preference profile where each of the three voters behave rationally. Then the electoral decision, the **social choice**, should be a rational (transitive) ranking of X,Y and Z.

Arrow's Impossibility Theorem 4

So Arrow proposes that there should be a **social welfare function (SWF)** which assigns to each rational preference profile a rational social preference profile.

	X vs Y								Y vs Z								Z vs X							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<
					>	?					>			?				<					?	

2. The SWF is to be specified by giving a value, < or >, to each column of the table. Thus, a valid choice of three columns (rational preference profile) will be assigned an election result (transitive pairwise comparison of X, Y and Z). In particular, **Arrow's 1st condition**, the SWF should not change for a particular column regardless of how the other columns of a preference profile change. This is **Independence of Irrelevant Alternatives (IIA).**

Economists IIA joke: Kenneth Arrow goes to a restaurant: what dessert would sir like, ice cream or apple pie. He'll take the apple pie. Waiter comes back: actually, there's also chocolate mousse. Arrow replies – oh, in that case I'll have the ice cream.

Arrow's Impossibility Theorem 5

So, we are adding a 4th row to our table: **S** records the values of our SWF.

	X vs Y								Y vs Z								Z vs X							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
A	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<
S	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**

Next, Arrow adds a final, very natural, condition: if everybody agrees on a comparison between two candidates, the SWF should reflect this choice.

3. **Unanimity** : if the values of A, B and C in any column are identical, then S must take this same value.

Arrow’s Impossibility Theorem 6

And this is what Arrow found to be the unavoidable consequence of his 4 conditions:

	X vs Y									Y vs Z									Z vs X							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**		**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**		**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
A	>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<
S	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**		**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**		**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**

Row **S** here is identical to row C. This means C has a certain absolute power! They can reverse their comparison of X vs Y being sure that **S**, the social choice, will switch to this new individual voter choice.
So this should be a 3rd condition for a fair democratic election.

4. **Non-dictatorship :** the SWF should not be identical to any one row of the table.

Arrow's Impossibility Theorem 7

And from these very plausible conditions, Arrow deduced his famous theorem:

There is no SWF which simultaneously satisfies Rationality, Unanimity, Independence of Irrelevant Alternatives and Non-dictatorship.

	X vs Y									Y vs Z									Z vs X							
	1	2	3	4	5	6	7	8		9	10	11	12	13	14	15	16		17	18	19	20	21	22	23	24
A	>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<		>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<		>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<		>	<	>	<	>	<	>	<
S	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**		**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**		**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**
					>							>								>						

Proof idea: (1) we can show that, if S agrees against the majority on a choice (comparison of two candidates), then the minority voter must be a dictator.
(2) Take a rational preference profile exhibiting Condorcet's paradox. Then either the SWF is non-transitive, or we are in case (1) which forces a dictatorship.

Proof of Arrow's Theorem 1

Arrow's Impossibility Theorem: *There is no SWF which satisfies U, IIA and ND.*

PROOF: The proof consists of two steps: in the first, I prove that if there is a disagreement on any given pairwise comparison, the SWF must agree with the majority. In the second step, I prove that this property implies the intransitivity of the SWF. Thus, the two steps jointly show that the axioms are inconsistent.

To prove the Theorem, extensive use will be made of the following table:

	x versus y							
	1	2	3	4	5	6	7	8
a	$\succ$	$\succ$	$\succ$	$\succ$	$\prec$	$\prec$	$\prec$	$\prec$
b	$\succ$	$\succ$	$\prec$	$\prec$	$\succ$	$\succ$	$\prec$	$\prec$
c	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	$\prec$
s	$\succ$							$\prec$
s		$\prec$	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	

	y versus z							
	9	10	11	12	13	14	15	16
a	$\succ$	$\succ$	$\succ$	$\succ$	$\prec$	$\prec$	$\prec$	$\prec$
b	$\succ$	$\succ$	$\prec$	$\prec$	$\succ$	$\succ$	$\prec$	$\prec$
c	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	$\prec$
s	$\succ$							$\prec$
s		$\prec$	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	

	z versus x							
	17	18	19	20	21	22	23	24
a	$\succ$	$\succ$	$\succ$	$\succ$	$\prec$	$\prec$	$\prec$	$\prec$
b	$\succ$	$\succ$	$\prec$	$\prec$	$\succ$	$\succ$	$\prec$	$\prec$
c	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	$\prec$
s	$\succ$							$\prec$
s		$\prec$	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	

Table 1

Step 1: Notice that in all the columns of the table which are not filled by U, we have a conflict between a majority of two individuals and a single dissenter. I first show that in the presence of such conflict, society cannot be indifferent. Secondly I show that if society sides with the single dissenting individual in a given case of disagreement, then this individual must be a dictator. Given that dictatorship is ruled out by axiom, this proves the step.

Take any case of disagreement of individual preferences on a given comparison, say, without loss of generality, column 2 in the table. Consider two alternative preference profiles: i) columns (2,9,23); and ii) columns (2,16,23). By IIA, social preference in the (x-y) and (z-x) comparisons must be the same in both cases. If in column 23 we have $x \sim_s z$, then $y \succ_s x \sim_s z$ in the first case and $z \sim_s x \succ_s y$ in the second. If $x \succ_s z$, transitivity forces $x \succ_s y$. If $z \succ_s x$, transitivity forces $y \succ_s x$. Therefore, in column 2 we must have either $x \succ_s y$ or $y \succ_s x$ (i.e. x cannot be socially indifferent to y).

Proof of Arrow's Theorem 2

Suppose then that $y \succ_s x$. Consider the preference profiles (2,16,19), (2,16,21) and (2,16,23). By assumption, $y \succ_s x$, and by U, $z \succ_s y$. Thus, by transitivity, $z \succ_s x$ in columns 19, 21 and 23 of row 5. Using the same reasoning, we choose the preference profiles: i) (2,11,24), (2,13,24), (2,15,24); ii) (1,15,18), (1,15,20), (1,15,22); iii) (1,10,23), (1,12,23), (1,14,23); iv) (3,16,18), (5,16,18), (7,16,18); v) (4,9,23), (6,9,23) to fill the other entries in that row.[1] But rows 4 and 5 jointly imply that social preference is identical to row 3, that is, c is a dictator, and the step is proved.[2]

	X vs Y								Y vs Z								Z vs X							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<
S	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**

	X vs Y								Y vs Z								Z vs X							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<	>	>	>	>	<	<	<	<
B	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<	>	>	<	<
C	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<	>	<
S	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**	**>**	**<**