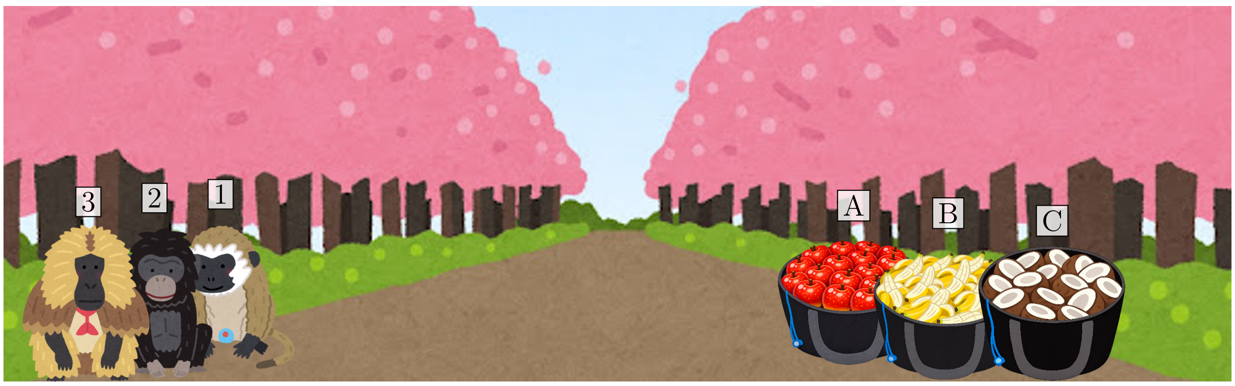
MAT120: Lecture 17 Handout

Mathematics and Social Choice

Last lecture we studied mathematics and money, focusing on growth, interest, expected value, and risk. In this lecture we move further into the "mathematics in society" block by asking how mathematics can be used to study collective decision-making. The central theme is *social choice theory*: how should a group combine many individual preferences into one social outcome? We will introduce some basic terminology, study several voting rules, look at two competing ideas of fairness, then finish with the spoiler effect and Condorcet cycles. The deeper lesson is that even something as familiar as "taking a vote" is subjective and hides some interesting mathematics.

1 What is Social Choice?

Suppose that three monkeys are walking through the forest, and they find three different bags of food: one bag is full of apples, one bag is full of bananas and one bag is full of coconuts. Suppose that our monkeys want to steal these fruits, but the bags are big and heavy, and therefore it takes three monkeys to pick up a single bag. With this in mind, the monkeys decide that they can only steal one of the three bags of fruit.



Of course, each monkey has their own preferences: perhaps Monkey 1 prefers bananas over coconuts and apples, or perhaps Monkey 3 really likes apples. The group must somehow decide together how to turn these preferences into a single, collective decision. How should they do this? At first this sounds simple: the monkeys should simply take a vote, count up the answers and pick a winner. However, we then run into the issue of *how* the votes should be counted.

This is the basic problem of *social choice theory*: how do we make a group decision based on a collection of individual preferences?

Of course, this is an extremely important question: beyond opportunistically kleptomaniacal primates, we humans mostly organise our societies democratically – all eligible members of society have an equal say in choosing a political leader. In a democratic setting like this, it would be nice to know whether or not our vote-counting system was fair, reasonable and mathematically well-behaved.

1.1 Normative versus empirical questions

It should be noted that social choice theory is mainly a *normative* subject, not an *empirical* one. That means it is trying to answer the question "what *should* we do?", rather than merely asking "what do we do?". This is Hume's famous is/ought distinction – some statements are simply descriptive in that they try to describe the world as it *is*, and some statements try to describe the world as it *ought to be*.

Normative and empirical claims

- A **normative** claim asks what the right thing to do is.
- An **empirical** claim asks what actually happens in the real world.

Of course, real elections involve psychology, history, media, institutions, cultural movements and strategy. But, before discussing those complications, it is useful to isolate the basic mathematical issue: once the voters have reported their preferences, what rule should transform those preferences into an outcome? This is what we will discuss today.

1.2 Basic terminology

As a matter of fact, there are many different types of voting rules. Before discussing these, we first need to introduce some new language.

Basic terminology of Social Choice Theory

- **Voters:** the group providing the input.
- **Options:** the possible outcomes or candidates.
- **Preference ranking:** one voter's ordering of the options from best to worst.
- **Preference profile:** the full list of all voters' rankings.
- **Social choice rule:** a method that takes a preference profile and produces an outcome.
- **Outcome:** the decision produced by the voting process.

In our example of the three monkeys and the bags of fruit:

- the **voters** are the monkeys,
- the **options** are the three bags of fruit,
- a **preference ranking** would be an individual monkey's preferred ranking of apples, bananas and coconuts,
- a **preference profile** would be a collection of three preference rankings (one for each monkey),
- a **social choice rule** would be a particular way in which the monkeys decided how to turn the collection of their individual preferences into a single decision, and
- the **outcome** would be the bag of fruit that they decide to steal.

1.3 Preference Profiles

A preference profile is usually written in a table, with one column for each voter and one row for each rank. Usually the top row shows the most preferred option, and the bottom row shows the least preferred option for each voter.

Using our three monkeys, one example of a preference profile is below:

	Monkey 1	**Monkey 2**	**Monkey 3**
1st choice:	A	A	B
2nd choice:	B	C	A
3rd choice:	C	B	C

It is also useful to introduce some mathematical notation. We will use the symbol $>$ to represent preference, so that $A > B$ is read "A is preferred to B". This way, we can also write a preference ranking mathematically, by taking the set of options and lining them up in order of preference, from left to right. For example, the preference profile written in the table above can also be read as:

- Monkey 1: $A > B > C$,
- Monkey 2: $A > C > B$,
- Monkey 3: $B > A > C$.

Observe that these are just the columns of the table above, written using $>$ instead.

2 Voting Rules

A central observation in social choice theory is that the same preference profile can produce totally different winners, depending on which voting rule we impose. So, the outcome of an election is not determined only by what the voters think, but also by the way in which we choose to count votes.

Important point

Different social choice rules can produce different winners, even when the voters' preferences are exactly the same.

Since we only have one lecture to discuss social choice theory, we will only introduce four basic social choice rules. These are:

1. Plurality,
2. Runoff,
3. the Borda count, and
4. the Condorcet count.

2.1 Plurality rule

Under the *plurality rule*, each voter chooses only one option and the winner is simply the option with the most first-place votes. This is the simplest counting rule, and it is widely used in practice.

Plurality intuition. Plurality looks only at the top row of the preference profile. Whoever appears most often in first place wins.

For example, consider the preference profile above. Here, plurality rule only cares about the first row, so we simply look at the top row and see what wins:

	Monkey 1	Monkey 2	Monkey 3
1st choice:	A	A	B
2nd choice:	B	C	A
3rd choice:	C	B	C

← A wins

Plurality is simple, but that simplicity comes at a cost: it ignores second-place and third-place information completely. An option that many people find "acceptable" but few people rank first can easily lose.

2.2 Runoff

A runoff rule tries to refine the plurality idea. At this introductory level, we can think of it as a second vote among the leading options from the first round. This can happen when the first round does not produce a decisive winner. This is meant to reduce the problem that plurality only uses a single layer of information.

The exact details of runoff systems vary in real elections. The main idea for us is simply this: if the first count does not settle matters decisively, the leading options are compared again in a more focused second stage.

2.3 Borda count

A *Borda count* uses the full preference profile. Instead of asking only "who is first?", it awards points according to position in each ranking. With three options, a natural scoring system is:

$$\text{1st place } = 3, \qquad \text{2nd place } = 2, \qquad \text{3rd place } = 1.$$

Then we add the points across all voters, and the option with the highest total wins.

Borda count procedure

If there are m options, assign points by rank, add the points for each option across all voters, and pick the option with the largest total.

As an example, consider again the preference profile from Section 2.1. Using the scoring rule $3, 2, 1$, we see that we obtain:

- Monkey 1 gives $A = 3$, $B = 2$, $C = 1$;
- Monkey 2 gives $A = 3$, $C = 2$, $B = 1$;
- Monkey 3 gives $B = 3$, $A = 2$, $C = 1$.

So the totals are

$$A = 3 + 3 + 2 = 8, \qquad B = 2 + 1 + 3 = 6, \qquad C = 1 + 2 + 1 = 4.$$

Hence the Borda winner is A. It is perhaps easiest to work this all out using the preference table from before. Here, the above reasoning can be summarized as:

	Monkey 1	Monkey 2	Monkey 3
1st choice	$A = 3$	$A = 3$	$B = 3$
2nd choice	$B = 2$	$C = 2$	$A = 2$
3rd choice	$C = 1$	$B = 1$	$C = 1$

$A : 3 + 3 + 2 =$ 8 points ⟵ A wins
$B : 2 + 1 + 3 =$ 6 points
$C : 1 + 2 + 1 =$ 4 points

2.3.1 Another Borda example

Now consider the profile

	Monkey 1	Monkey 2	Monkey 3
1st choice	B	C	B
2nd choice	C	A	C
3rd choice	A	B	A

Here, the Borda totals are

$$A = 1 + 2 + 1 = 4, \qquad B = 3 + 1 + 3 = 7, \qquad C = 2 + 3 + 2 = 7.$$

So in this case the Borda count gives a *tie* between B and C. In summary, we have:

	Monkey 1	Monkey 2	Monkey 3
1st choice	$B = 3$	$C = 3$	$B = 3$
2nd choice	$C = 2$	$A = 2$	$C = 2$
3rd choice	$A = 1$	$B = 1$	$A = 1$

$A : 1 + 2 + 1 =$ 4 points
$B : 3 + 1 + 3 =$ 7 points
$C : 2 + 3 + 2 =$ 7 points ⟵ a tie

2.4 Condorcet count

A *Condorcet count* is similar to the Borda count, in that it also uses the full preference profile. However, a Condorcet count reads the information in a different way: instead of assigning points, it compares *pairs of options* to each other. For example, if there were three options A, B and C, then we would need to compare A against B, then A against C, and then B against C. For each pair, we count how many voters prefer one option to the other. The Condorcet count uses *this* information to decide who should win.

Condorcet idea

An option is a **Condorcet winner** if it defeats every other option in one-on-one comparisons.

Returning to our monkeys, consider again the preference profile:

- Monkey 1: $A > B > C$,
- Monkey 2: $A > C > B$,

- Monkey 3: $B > A > C$.

Since there are only 3 options, there are 3 pairwise comparisons that we need to make.

1. A versus B: two voters prefer A to B, so A beats B.
2. A versus C: all three voters prefer A to C, so A beats C.
3. B versus C: two voters prefer B to C, so B beats C.

Since A beats both B and C one-on-one, A is the Condorcet winner.

In terms of the preference table, we can determine the Condorcet winner by looking at the pairs of options and counting which option wins in each column. For example, the comparison of Apples to Bananas written above can be visualized as:

	Monkey 1	**Monkey 2**	**Monkey 3**
1st choice:	A	A	B
2nd choice:	B	C	A
3rd choice:	C	B	C

$A > B$ twice $\longleftarrow$ A beats B
$B > A$ once

For completeness, the tabular forms of comparisons (2) and (3) are also depicted below:

	Monkey 1	**Monkey 2**	**Monkey 3**
1st choice:	A	A	B
2nd choice:	B	C	A
3rd choice:	C	B	C

$A > C$ thrice $\longleftarrow$ A beats C
$C > A$ never

	Monkey 1	**Monkey 2**	**Monkey 3**
1st choice:	A	A	B
2nd choice:	B	C	A
3rd choice:	C	B	C

$B > C$ twice $\longleftarrow$ B beats C
$C > B$ once

As you can see, based on the three pairwise comparisons, A simultaneously beats both B and C in a head-to-head comparison, and therefore it is deemed to be the Condorcet winner.

2.5 Borda and Condorcet can disagree

The Borda count and the Condorcet count are both trying to use more information than plurality, but they need not agree. In fact, with only three options and three voters, it is not possible to obtain different unique winners under the Borda and Condorcet methods, although ties can still occur. However, if we add two more voters, then we may create such a profile. Consider the following:

	1	**2**	**3**	**4**	**5**
1st choice	A	B	A	B	A
2nd choice	B	C	B	C	B
3rd choice	C	A	C	A	C

We can check both rules explicitly. We start with the Borda count. Observe that the preference profile is:

- Voter 1: $A > B > C$,
- Voter 2: $B > C > A$,
- Voter 3: $A > B > C$,
- Voter 4: $B > C > A$,
- Voter 5: $A > B > C$.

So, the Borda scores are:

$$A = 3 + 1 + 3 + 1 + 3 = 11,$$

$$B = 2 + 3 + 2 + 3 + 2 = 12,$$

$$C = 1 + 2 + 1 + 2 + 1 = 7.$$

Therefore, the Borda count says that B is the winner.

Next we compute the Condorcet count. The pairwise comparisons are:

1. A **versus** B**:** Voters 1, 3 and 5 prefer A to B, while Voters 2 and 4 prefer B to A. So, A beats B.
2. A **versus** C**:** Voters 1, 3 and 5 prefer A to C, while Voters 2 and 4 prefer C to A. So, A beats C.
3. B **versus** C**:** every voter prefers B to C, so B beats C.

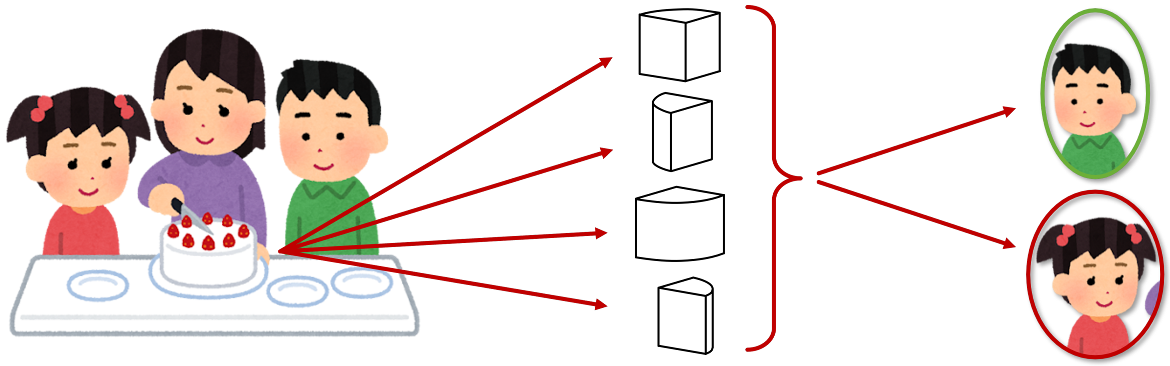
Since A beats both B and C in head-to-head comparisons, the Condorcet winner is A.

For this preference profile, the Borda count chooses B, but the Condorcet count chooses A. This shows that even when two voting rules use the full preference profile, they can still produce different winners.

3 Fairness and Utility

Voting is not the only setting in which a group must make a collective decision. Another important question is how to divide resources fairly – a process known as "fair division". In the interest of time, we will only introduce the idea through a single example.

Suppose that a mother is trying to divide a birthday cake between her two children. She cuts the cake into four slices and would like to give them to her children:



How should she assign these slices of cake?

3.1 Utility Functions

Suppose that we label the four slices of cake A, B, C, and D. One approach to the assignment of cake slices would be to simply ask the children which slices they prefer, and then to figure out how to give the slices to each child in a fair way. Each slice of cake may have a different value for each child, so we may ask the children to give each cake slice a score out of 10, where a higher number corresponds to a more valuable slice of cake. For example, perhaps the girl dislikes icing, so she would value a slice with lots of icing as a 3. Perhaps the boy is feeling particularly hungry, so he values the bigger slice of cake as an 8.

Suppose that the children value each slice of cake as follows:

	A	B	C	D
Boy	4	8	5	10
Girl	0	4	7	8

Each child's assignment of values is known as a *utility function*. It tells us how much each person values each possible piece. Once we have this data in hand, we can begin to study fairness from a numerical perspective.

3.2 Assignments as subsets

Any allocation of cake slices can be described using set theory. Here, we can take any allocation of cake slices that the boy or girl gets, and identify it with a subset of $\{A, B, C, D\}$. Of course, one half of this assignment is fixed, because if the boy gets assigned some subset, then the girl will be assigned its complement. For example:

- if the boy gets slices A, B and C, then the girl gets the slice D. The assignment gives the subset $\{A, B, C\}$ to the boy, and the girl gets given the subset $\{D\}$;
- if the boy gets $\{A, C\}$, then the girl gets $\{B, D\}$;
- if the boy gets all four slices, then the girl gets $\varnothing$.

In principle, the boy's assignment can be any subset of $\{A, B, C, D\}$. So, the set of all possible assignments is given by the powerset

$$\mathcal{P}(\{A, B, C, D\}).$$

Since a set of size 4 has $2^4 = 16$ subsets, there are 16 possible assignments.

Why the powerset appears

Choosing an allocation is equivalent to choosing which subset of slices the boy receives. Once that subset is chosen, the girl's share is automatically its complement. So the number of possible allocations is the size of the powerset, namely $2^4 = 16$.

3.3 Counting the value of an assignment

To evaluate an assignment, we will simply add the relevant utilities contained in that subset.[1] For instance, if the boy gets $\{B, C\}$ and the girl gets $\{A, D\}$, then:

$$\text{value for the boy} = 8 + 5 = 13, \qquad \text{value for the girl} = 0 + 8 = 8.$$

The following table records all possible allocations. For future convenience, we will also add two extra columns to the end of the table.

Boy gets	Girl gets	Boy's value	Girl's value	Sum	Minimum
$\varnothing$	$\{A, B, C, D\}$	0	19	19	0
$\{A\}$	$\{B, C, D\}$	4	19	23	4
$\{B\}$	$\{A, C, D\}$	8	15	23	8
$\{C\}$	$\{A, B, D\}$	5	12	17	5
$\{D\}$	$\{A, B, C\}$	10	11	21	10
$\{A, B\}$	$\{C, D\}$	12	15	27	12
$\{A, C\}$	$\{B, D\}$	9	12	21	9
$\{A, D\}$	$\{B, C\}$	14	11	25	11
$\{B, C\}$	$\{A, D\}$	13	8	21	8
$\{B, D\}$	$\{A, C\}$	18	7	25	7
$\{C, D\}$	$\{A, B\}$	15	4	19	4
$\{A, B, C\}$	$\{D\}$	17	8	25	8
$\{A, B, D\}$	$\{C\}$	22	7	29	7
$\{A, C, D\}$	$\{B\}$	19	4	23	4
$\{B, C, D\}$	$\{A\}$	23	0	23	0
$\{A, B, C, D\}$	$\varnothing$	27	0	27	0

3.4 Two theories of fairness

Once the utilities have been calculated, different ethical principles can be imposed. We will only focus on two for now.

[1] It should be noted that this is an assumption of our approach. It does not need to be the case, in general, that the value of an assignment $\{A, B\}$ must equal the sum of the individual values of A and B. We can imagine situations in which this might *not* be the case, for example, the values of eating chicken and eating jam in a sandwich might be quite high, but the value of having a chicken and jam sandwich would be less.

Two theories of fairness

- **Utilitarianism:** choose the assignment that maximises the *sum* of all utilities.
- **Egalitarianism:** choose the assignment that maximises the *minimum* utility across the people involved.

These two rules are not the same. Utilitarianism tries to find an allocation that would make the total value as large as possible, whereas egalitarianism tries to find an allocation that would protect the worst-off person as much as possible.

3.4.1 Utilitarian conclusion

Looking at the sum column in the table above, the largest total value is 29, which is achieved when

$$\text{Boy gets } \{A, B, D\}, \qquad \text{Girl gets } \{C\}.$$

So the utilitarian selects that assignment.

3.4.2 Egalitarian conclusion

Looking at the minimum column in the table above, the largest minimum is 12, which is achieved when

$$\text{Boy gets } \{A, B\}, \qquad \text{Girl gets } \{C, D\}.$$

So the egalitarian selects a different assignment.

Fairness is principle-dependent

For this example:

- **Utilitarianism** chooses $\{A, B, D\}$ for the boy and $\{C\}$ for the girl.
- **Egalitarianism** chooses $\{A, B\}$ for the boy and $\{C, D\}$ for the girl.

Therefore "the fairest assignment" is not absolute or necessarily objective. It depends on the principle of fairness that we choose to adopt.

4 The Spoiler Effect

The *spoiler effect* is one of the most famous problems associated with plurality-style voting. It occurs when a third candidate changes the final outcome, even though that candidate had no realistic chance of winning.

The most well-known example comes from the 2000 US election, for the state of Florida. In this situation, there were 3 main candidates: George W. Bush, Al Gore and Ralph Nader. The Florida vote totals for these three candidates were:

	Bush	**Gore**	**Nader**
Votes	2,912,790	2,912,253	97,421

Therefore Bush won the state of Florida by a tiny margin of 537 votes.

Nader, being a Green Party candidate, was further to the political left than Al Gore was. Therefore, if given the choice, many of the people who voted for Nader would probably choose Gore over Bush in a head-to-head comparison. This might plausibly have led to Gore winning Florida. However, the *mere existence* of Nader as an option may have drawn these votes away from Gore, and therefore in reality he lost.

Spoiler effect

A **spoiler effect** occurs when a candidate with little chance of winning changes the final outcome by pulling votes away from a nearby candidate.

In other words, adding an "irrelevant" option may fail to be irrelevant. This is one reason social choice theory cares not only about winners, but also about how sensitive a rule is to the presence of extra options.

5 Condorcet Cycles

One might hope that using a pairwise majority comparison is always enough to settle an election cleanly. Condorcet cycles show that this hope is false. Consider the preference profile:

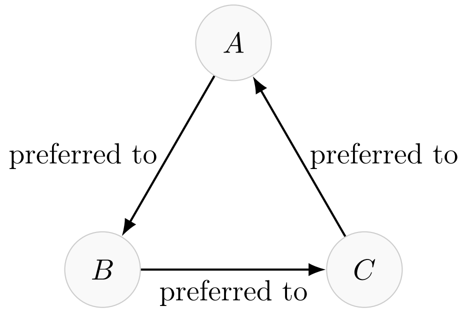
	Voter 1	**Voter 2**	**Voter 3**
1st choice	A	C	B
2nd choice	B	A	C
3rd choice	C	B	A

For this profile, there is no majority winner, and plurality gives a 3-way tie. A runoff system would therefore need some additional tie-breaking rule or more detailed specification. Moreover, the Borda count gives equal scores to all three options, so there is a tie. Most interestingly, the Condorcet count gives the following pairwise comparisons:

The lecture summarises the consequences as follows:

- $A > B$ twice and $B > A$ once, so A beats B,
- $C > A$ twice and $A > C$ once, so C beats A, and
- $B > C$ twice and $C > B$ once, so B beats C.

Thus we have created a "preference loop" in which the group prefers A to B and prefers B to C, yet also prefers C to A:



Condorcet cycle

A **Condorcet cycle** happens when majority preference becomes cyclic:

$$A > B, \qquad B > C, \qquad \text{and} \qquad C > A.$$

In this situation there is no Condorcet winner.

This is a very interesting result: we have just demonstrated that even in very simple situations, individual voters may each have perfectly consistent rankings, but the group preference can still fail to be globally consistent.

6 Exercises

Exercise 1: voting rules

Consider the preference profile

	Voter 1	**Voter 2**	**Voter 3**
1st choice	B	C	B
2nd choice	C	A	C
3rd choice	A	B	A

(a) Who wins under plurality?

(b) Compute the Borda totals using the scoring rule $3, 2, 1$.

(c) Determine the Condorcet winner by pairwise comparison.

Solution

(a) Under plurality, only first-place votes matter. B appears twice in first place and C once, so plurality chooses B.

(b) The Borda totals are

$$A = 1 + 2 + 1 = 4, \qquad B = 3 + 1 + 3 = 7, \qquad C = 2 + 3 + 2 = 7.$$

So Borda gives a tie between B and C.

(c) Pairwise comparisons give:

- B beats A by 2 votes to 1,
- C beats A by 3 votes to 0,
- B beats C by 2 votes to 1.

Therefore B beats both other options one-on-one, so the Condorcet winner is B.

Exercise 2: fairness and utility

Use the utility table

	A	B	C	D
Boy	4	8	5	10
Girl	0	4	7	8

Suppose the boy gets $\{A, D\}$ and the girl gets $\{B, C\}$.

(a) Compute the boy's value.

(b) Compute the girl's value.

(c) Compute the utilitarian total and the egalitarian minimum.

Solution

The boy's value is

$$4 + 10 = 14.$$

The girl's value is

$$4 + 7 = 11.$$

So the utilitarian total is

$$14 + 11 = 25,$$

and the egalitarian minimum is

$$\min(14, 11) = 11.$$