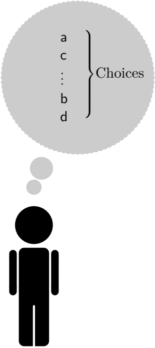
(Formalization) of Social Choice Theory

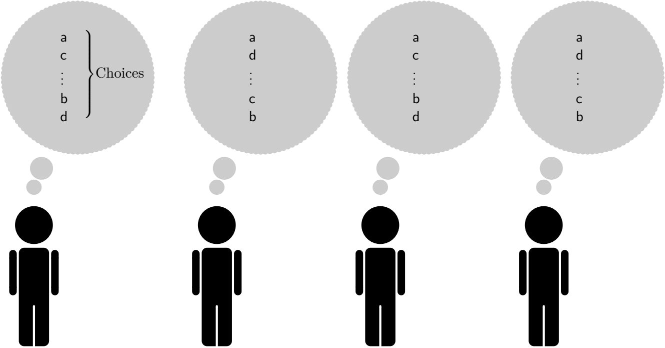
Julian Parsert
supervised by Cezary Kaliszyk

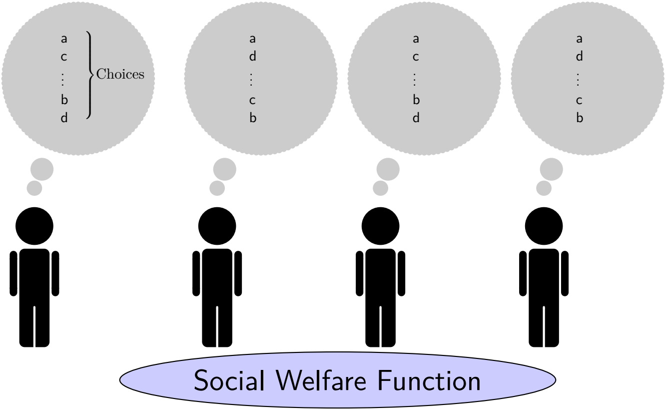
Department of Computer Science

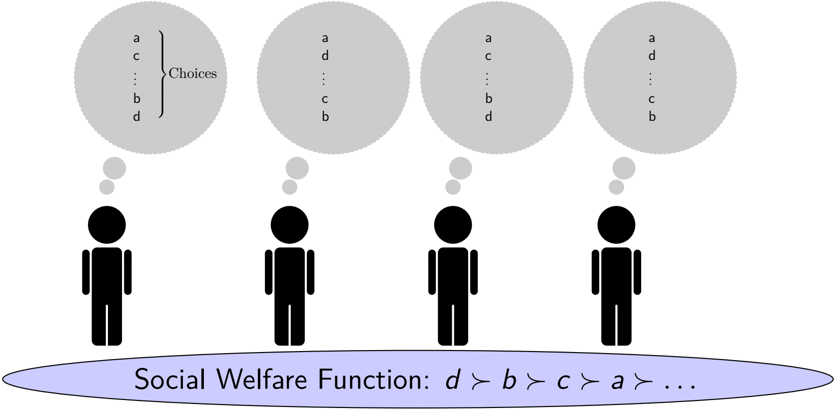
Innsbruck November 23, 2017

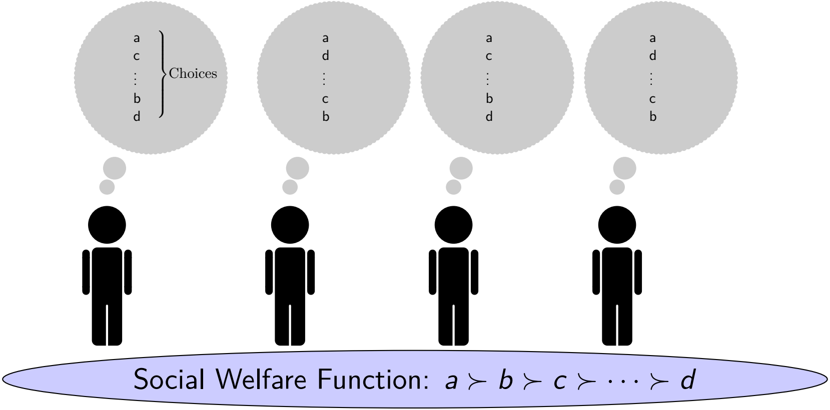
Contents

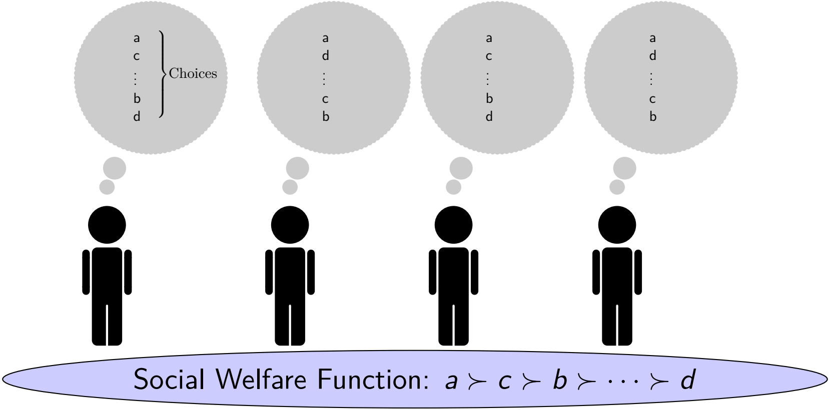

a
c
⋮
b
d
Choices


a
c
⋮
b
d
Choices
a
d
⋮
c
b
a
c
⋮
b
d
a
d
⋮
c
b


a
c
⋮
b
d
Choices
a
d
⋮
c
b
a
c
⋮
b
d
a
d
⋮
c
b
Social Welfare Function


a
c
⋮
b
d
Choices
a
d
⋮
c
b
a
c
⋮
b
d
a
d
⋮
c
b
Social Welfare Function: $d \succ b \succ c \succ a \succ \ldots$


a
c
⋮
b
d
Choices
a
d
⋮
c
b
a
c
⋮
b
d
a
d
⋮
c
b
Social Welfare Function: $a \succ b \succ c \succ \cdots \succ d$


a
c
⋮
b
d
Choices
a
d
⋮
c
b
a
c
⋮
b
d
a
d
⋮
c
b
Social Welfare Function: $a \succ c \succ b \succ \cdots \succ d$

Definition (Preference Relation over $\mathcal{C}$)

binary relation, $\succeq$ over $\mathcal{C}$
transitive, reflexive

Definition (Rational Preference Relation over $\mathcal{C}$)

binary relation, $\succeq$ over $\mathcal{C}$
transitive, reflexive, total

Definition (Rational Preference Relation over $\mathcal{C}$)

binary relation, $\succeq$ over $\mathcal{C}$
transitive, reflexive, total
$x \succ y \equiv x \succeq y \land \neg y \succeq x$

Definition (Rational Preference Relation over $\mathcal{C}$)

binary relation, $\succeq$ over $\mathcal{C}$
transitive, reflexive, total
$x \succ y \equiv x \succeq y \land \neg y \succeq x$

Example (Preference Relation over a, b, c for Alice)

$$P_{Alice} := \{(b, c), (b, a), (c, a)\} \equiv \begin{array}{c} P_{Alice} \\ \hline \text{b} \\ \text{c} \\ \text{a} \end{array} \equiv b \succ_{Alice} c \succ_{Alice} a$$

Social Welfare Functions

The Setup

- A non empty set of n Individuals I
- A set of choices, $\mathcal{C}$ with $|\mathcal{C}| \geq 3$
- $\mathcal{P}$ is the set of all possible preference relations over $\mathcal{C}$

Social Welfare Functions

The Setup

- A non empty set of n Individuals I
- A set of choices, $\mathcal{C}$ with $|\mathcal{C}| \geq 3$
- $\mathcal{P}$ is the set of all possible preference relations over $\mathcal{C}$

Notation

$x \succeq_{P_i} y \equiv (x, y) \in P_i$
$x \succeq_i y \equiv x \succeq_{P_i} y$
$x \succ_i y \equiv x \succeq_i y \wedge \neg y \succeq_i x$
$P_i \in \mathcal{P}$ is the set of pairs describing preferences of agent i

Social Welfare Functions

The Setup

- A non empty set of n Individuals I
- A set of choices, $\mathcal{C}$ with $|\mathcal{C}| \geq 3$
- $\mathcal{P}$ is the set of all possible preference relations over $\mathcal{C}$

Notation

$x \succeq_{P_i} y \equiv (x, y) \in P_i$
$x \succeq_i y \equiv x \succeq_{P_i} y$
$x \succ_i y \equiv x \succeq_i y \wedge \neg y \succeq_i x$
$P_i \in \mathcal{P}$ is the set of pairs describing preferences of agent i

Assumption

Agents preferences are strict: $\forall i \in I.\ x \succeq_i y \rightarrow x \succ_i y$

Social Welfare Function

Definition (Social Welfare Function (SWF))

A social welfare function is a mapping $\mathcal{F} : \mathcal{P}^n \to \mathcal{P}$

Social Welfare Function

Definition (Social Welfare Function (SWF))

A social welfare function is a mapping $\mathcal{F} : \mathcal{P}^n \to \mathcal{P}$

Notation

$x \succeq y \equiv x \succeq_{SWF(P_1,\dots,P_n)} y \equiv (x, y) \in \mathcal{F}(P_1, \dots, P_n)$
$x \succ y \equiv x \succeq y \wedge \neg y \succeq x$
Preference Profile P where $P := (P_1, \dots, P_n) \in \mathcal{P}^n$

Definition (Dictatorial SWF)

- A Social Welfare Function $\mathcal{F}$ is dictatorial if there exists individual $i \in I$ such that $\forall xy \in \mathcal{C}.\ x \succ_i y \rightarrow x \succ y$
- $\mathcal{F}$ is dictatorial for i if $\mathcal{F}$ is a projection π_i^n.

Definition (Dictatorial SWF)

- A Social Welfare Function $\mathcal{F}$ is dictatorial if there exists individual $i \in I$ such that $\forall xy \in \mathcal{C}.\ x \succ_i y \rightarrow x \succ y$
- $\mathcal{F}$ is dictatorial for i if $\mathcal{F}$ is a projection π_i^n.

1	...	i	...	n	$\mathcal{F}$
...	...	b	⋮	...	b
⋮	⋮	X	⋮	⋮	X
...	...	a	...	...	a
⋮	⋮	Y	⋮	⋮	Y
...	...	c	...	...	c
⋮	...	Z	...	...	Z

Pivotal Individual

Definition (Decisive/Pivotal)

- An agent i is pivotal for $a, b \in \mathcal{C}$ at $P_i, \dots P_n$ if by changing the relation between a and b the SWF changes this relation as well.
- $a \succ_i b \iff a \succ_{SWF(P_1,\dots,P_i,\dots,P_n)} b$

Unanimity

Unanimity

- If all individuals have the same relation between two alternatives a and b, the SWF has to reflect that relation.
- $\forall a\ b \in \mathcal{C}.\ (\forall i \in I.\ a \succeq_i b) \rightarrow a \succeq b$

Unanimity

Unanimity

- If all individuals have the same relation between two alternatives a and b, the SWF has to reflect that relation.
- $\forall a\ b \in \mathcal{C}.\ (\forall i \in I.\ a \succeq_i b) \rightarrow a \succeq b$

1	...	...	...	n	$\mathcal{F}$
a	...	...	...	a	⋮
⋮	a	⋮	⋮	⋮	a
...	b	...	a	...	⋮
⋮	⋮	⋮	⋮	⋮	⋮
b	...	a	b	...	b
⋮	...	b	...	b	⋮

Independence of Irrelevant Alternatives

Independence of Irrelevant Alternatives (IIA)

- Social preference between two alternatives a and b only depends on the voters' preferences between a and b.

Independence of Irrelevant Alternatives

Independence of Irrelevant Alternatives (IIA)

- Social preference between two alternatives *a* and *b* only depends on the voters' preferences between *a* and *b*.
- Given preference relations $P_1, \ldots, P_n \in \mathcal{P}$ and $Q_1, \ldots, Q_n \in \mathcal{P}$ then

$$\forall a\ b \in \mathcal{C}.\ (\forall i \in I.\ a \succ_{P_i} b \iff a \succ_{Q_i} b)$$
$$\Longrightarrow$$
$$(a \succ_{SWF(P_1,\ldots,P_n)} b \iff a \succ_{SWF(Q_1,\ldots,Q_n)} b)$$

Independence of Irrelevant Alternatives cont.

1	...	...	...	n	$\mathcal{F}$
a	a	c	...	a	...
⋮	c	⋮	...	⋮	a
c	b	...	a	c	...
⋮	⋮	⋮	c	⋮	c
b	...	a	b	...	b
⋮	...	b	...	b	...

Independence of Irrelevant Alternatives cont.

1	...	...	...	n	$\mathcal{F}$
c	c	c	...	a	...
⋮	a	⋮	...	⋮	c
a	b	...	a	c	...
⋮	⋮	⋮	c	⋮	a
b	...	a	b	...	b
⋮	...	b	...	b	...

Independence of Irrelevant Alternatives cont.

1	...	...	...	n	$\mathcal{F}$
c	c	c	a	a	...
⋮	a	⋮	c	⋮	c
a	b	...	...	c	...
⋮	⋮	⋮	⋮	⋮	a
b	...	a	b	...	b
⋮	...	b	...	b	...

Preliminary Lemma

Lemma (Strict Neutrality Lemma)

Consider two pairs $u, v \in \mathcal{C}$ *and* $a, b \in \mathcal{C}$. *Suppose all* i *have the same relation between* u, v *and* a, b ($u \succ v \iff a \succ b$). *Then the SWF ranks (strictly)* u, v *the same as* a, b.

Theorem (Arrow's Impossibility Theorem)

Any Social Welfare Function that respects transitivity, independence of irrelevant alternatives (IIA) and unanimity is dictatorial.

Theorem (Arrow's Impossibility Theorem)

Any Social Welfare Function that respects transitivity, independence of irrelevant alternatives (IIA) and unanimity is dictatorial.

Example (Naive Implementation)

1. Take two alternatives x and y and see which receives more votes.
2. Repeat for every pair.

Theorem (Arrow's Impossibility Theorem)

Any Social Welfare Function that respects transitivity, independence of irrelevant alternatives (IIA) and unanimity is dictatorial.

Example (Naive Implementation)

1. Take two alternatives x and y and see which receives more votes.
2. Repeat for every pair.

Problem:

c	a	b
b	c	a
a	b	c

Theorem (Arrow's Impossibility Theorem)

Any Social Welfare Function that respects transitivity, independence of irrelevant alternatives (IIA) and unanimity is dictatorial.

Example (Naive Implementation)

1. Take two alternatives x and y and see which receives more votes.
2. Repeat for every pair.

Problem:

c	a	b	$c \succ b$
b	c	a	$b \succ a$
a	b	c	$a \succ c$
			$\lightning$

Examples cont.

Example (Allow Indifference)

c	a	b	$c \approx b$
b	c	a	$b \approx a$
a	b	c	

Examples cont.

Example (Allow Indifference)

c	a	a
b	c	b
a	b	c

Examples cont.

Example (Allow Indifference)

c	a	a	$a \succ b$
b	c	b	
a	b	c	

Examples cont.

Example (Allow Indifference)

c	a	a	$a \succ b$
b	c	b	$c \succ b$
a	b	c	

Examples cont.

Example (Allow Indifference)

c	a	a	$a \succ b$
b	c	b	$c \succ b$
a	b	c	$a \succ c$
			$\lightning$(IIA)

Proof

Theorem (Arrow's Impossibility Theorem)

Any Social Welfare Function that respects transitivity, independence of irrelevant alternatives (IIA) and unanimity is dictatorial.

Identify pivotal voter e

			$\mathcal{F}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\ldots$	$\ldots$	$\ldots$	b
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\ldots$	$\ldots$	$\ldots$	a
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Proof

Theorem (Arrow's Impossibility Theorem)

Any Social Welfare Function that respects transitivity, independence of irrelevant alternatives (IIA) and unanimity is dictatorial.

Identify pivotal voter e

1						n	$\mathcal{F}$
⋮	...	⋮	⋮	⋮	⋮	...	⋮
a	...	a	b	b	b	...	b
⋮	...	⋮	⋮	⋮	⋮	...	⋮
b	...	b	a	a	a	...	a
⋮	...	⋮	⋮	⋮	⋮	...	⋮

Proof

Theorem (Arrow's Impossibility Theorem)

Any Social Welfare Function that respects transitivity, independence of irrelevant alternatives (IIA) and unanimity is dictatorial.

Identify pivotal voter e

1						n	$\mathcal{F}$
⋮	…	⋮	⋮	⋮	⋮	…	⋮
a	…	a	a	b	b	…	b
⋮	…	⋮	⋮	⋮	⋮	…	⋮
b	…	b	b	a	a	…	a
⋮	…	⋮	⋮	⋮	⋮	…	⋮

Proof

Theorem (Arrow's Impossibility Theorem)

Any Social Welfare Function that respects transitivity, independence of irrelevant alternatives (IIA) and unanimity is dictatorial.

Identify pivotal voter e

1				e		n	$\mathcal{F}$
⋮	…	⋮	⋮	⋮	⋮	…	⋮
a	…	a	a	a	b	…	a
⋮	…	⋮	⋮	⋮	⋮	…	⋮
b	…	b	b	b	a	…	b
⋮	…	⋮	⋮	⋮	⋮	…	⋮

Proof cont.

Proof.

Show that pivotal voter e is a dictator ($\exists i.\ x \succ_i y \rightarrow x \succ y$)

□

Proof cont.

Proof.

Show that pivotal voter e is a dictator ($\exists i.\ x \succ_i y \rightarrow x \succ y$)

- Take arbitrary pair $x, y \in \mathcal{C}$ such that $x \succ_e y$

1			e			n	$\mathcal{F}$
⋮	…	⋮	⋮	⋮	…	⋮	
⋮	⋮	⋮	x	⋮	⋮	⋮	
⋮	…	⋮	⋮	⋮	…	⋮	
⋮	…	⋮	y	⋮	…	⋮	
⋮	…	⋮	⋮	⋮	…	⋮	

□

Proof cont.

Proof.

Show that pivotal voter e is a dictator $(\exists i.\ x \succ_i y \rightarrow x \succ y)$

- Take arbitrary pair $x, y \in \mathcal{C}$ such that $x \succ_e y$
- Take z such that $x \succ_e z \succ_e y$

1			e			n	$\mathcal{F}$
⋮	…	⋮	⋮	⋮	…	⋮	
⋮	⋮	⋮	x	⋮	⋮	⋮	
⋮	…	⋮	z	⋮	…	⋮	
⋮	…	⋮	y	⋮	…	⋮	
⋮	…	⋮	⋮	⋮	…	⋮	

□

Proof cont.

Proof.

Show that pivotal voter e is a dictator ($\exists i.\ x \succ_i y \to x \succ y$)

- Take arbitrary pair $x, y \in \mathcal{C}$ such that $x \succ_e y$
- Take z such that $x \succ_e z \succ_e y$

1			e			n	$\mathcal{F}$
z	…	z	⋮	⋮	…	⋮	
⋮	⋮	⋮	x	⋮	⋮	⋮	
⋮	…	⋮	z	⋮	…	⋮	
⋮	…	⋮	y	⋮	…	⋮	
⋮	…	⋮	⋮	z	…	z	

- $x \succ y$.

□

Proof cont.

Proof.

Show that pivotal voter e is a dictator ($\exists i.\ x \succ_i y \rightarrow x \succ y$)

- Take arbitrary pair $x, y \in \mathcal{C}$ such that $x \succ_e y$
- Take z such that $x \succ_e z \succ_e y$

1			e			n	$\mathcal{F}$
z	…	z	⋮	⋮	…	⋮	
⋮	⋮	⋮	x	⋮	⋮	⋮	$x \succ z$
⋮	…	⋮	z	⋮	…	⋮	
⋮	…	⋮	y	⋮	…	⋮	
⋮	…	⋮	⋮	z	…	z	

- $x \succ y$.

□

Proof cont.

Proof.

Show that pivotal voter e is a dictator $(\exists i.\ x \succ_i y \rightarrow x \succ y)$

- Take arbitrary pair $x, y \in \mathcal{C}$ such that $x \succ_e y$
- Take z such that $x \succ_e z \succ_e y$

1			e			n	$\mathcal{F}$
z	...	z	⋮	⋮	...	⋮	
x	x	x	x	x	x	x	$x \succ z$
⋮	...	⋮	z	⋮	...	⋮	
⋮	...	⋮	y	⋮	...	⋮	
⋮	...	⋮	⋮	z	...	z	

- $x \succ y$.

□

Proof cont.

Proof.

Show that pivotal voter e is a dictator $(\exists i.\ x \succ_i y \rightarrow x \succ y)$

- Take arbitrary pair $x, y \in \mathcal{C}$ such that $x \succ_e y$
- Take z such that $x \succ_e z \succ_e y$

1			e			n	$\mathcal{F}$
z	…	z	⋮	⋮	…	⋮	
⋮	⋮	⋮	x	⋮	⋮	⋮	$x \succ z$
⋮	…	⋮	z	⋮	…	⋮	$z \succ y$
⋮	…	⋮	y	⋮	…	⋮	
⋮	…	⋮	⋮	z	…	z	

- $x \succ y$.

□

Proof cont.

Proof.

Show that pivotal voter e is a dictator $(\exists i.\ x \succ_i y \rightarrow x \succ y)$

- Take arbitrary pair $x, y \in \mathcal{C}$ such that $x \succ_e y$
- Take z such that $x \succ_e z \succ_e y$

1			e			n	$\mathcal{F}$
z	...	z	⋮	⋮	...	⋮	
⋮	⋮	⋮	x	⋮	⋮	⋮	$x \succ z$
⋮	...	⋮	z	⋮	...	⋮	$z \succ y$
⋮	...	⋮	y	⋮	...	⋮	$x \succ y$
⋮	...	⋮	⋮	z	...	z	

- $x \succ y$.

□

Proof cont.

Proof.

Show that pivotal voter e is a dictator ($\exists i.\ x \succ_i y \rightarrow x \succ y$)

- Take arbitrary pair $x, y \in \mathcal{C}$ such that $x \succ_e y$
- Take z such that $x \succ_e z \succ_e y$

1			e			n	$\mathcal{F}$
z	...	z	⋮	⋮	...	⋮	
⋮	⋮	⋮	x	⋮	⋮	⋮	$x \succ z$
⋮	...	⋮	z	⋮	...	⋮	$z \succ y$
⋮	...	⋮	y	⋮	...	⋮	$x \succ y$
⋮	...	⋮	⋮	z	...	z	

- $x \succ y$. We identified an individual e who is a dictator

□

Formalization of Arrow's Theorem

Formalizations

- Mizar formalization by Freek Wiedijk
- Isabelle formalization by Nipkow
- Isabelle formalization by Peter Gammie

Further formalized results

Definition (Social Choice Function (SCF))

A social choice function is a mapping $\mathcal{G} : \mathcal{P} \times \cdots \times \mathcal{P} \to \mathcal{C}$

Theorem (Gibbard-Satterthwaite)

If $\mathcal{G}$ is incentive compatible and $|\mathcal{C}| \geq 3$ then $\mathcal{G}$ has a dictator.

Theorem (May's Theorem)

Majority rule is the only social choice function where $|\mathcal{C}| = 2$ and n is odd that is (1) anonymous (2) neutral (3) monotone.

Thanks for your attention

Lemma (Strict Neutrality Lemma)

Consider two pairs $u, v \in \mathcal{C}$ and $a, b \in \mathcal{C}$. Suppose all i have the same relation between u, v and a, b ($u \succ v \iff a \succ b$). Then the SWF ranks (strictly) u, v the same as a, b.

- assume u, v and a, b are not identical and $a \succeq b$
- Case $\{u, v, a, b\}$ are distinct

u	u	u	b	b	$a \succeq b$ by asm.
a	a	a	v	v	$u \succ a \wedge b \succ v$ by (U)
b	b	b	u	u	$u \succ v$
v	v	v	a	a	

- to show: $u \succ v$

□

Example (21 people rank 3 alternatives A, B and C)

1.	A	A	B	C	
2.	B	C	C	B	?
3.	C	B	A	A	
#votes	1	7	7	6	(21)