Lecture 15

Wednesday, 19 October 2022 12:13 AM

Mechanism Design

So far, game was given to us.
Now, we want to design games that have "good" properties

There are a set of alternatives $A = \{a_1, \dots, a_n\}$
(also called "outcomes")
Each agent i has value $v_i(a)$ for each alternative a (cardinal)
OR has a total order π_i over the alternatives (ordinal)

$\pi_i(a) > \pi_i(b) \Rightarrow$ agent i prefers alternative a to b

Mechanism takes as input some information abt. v_i / π_i from each agent i, picks an alternative $a^* \in A$
(in mechanisms with money, mechanism may also take / give payments — aka transfers — from/to agents)

Examples

(i) In an auction of a single item w/ n bidders, there are $n+1$ alternatives: which bidder the item goes to, or nobody.
Each bidder i submits v_i, its value for the alternative when the item is assigned to itself.
Mechanism picks an alternative, takes payment from winning bidder

(ii) In an election w/ n voters & m candidates, there are m alternatives (or perhaps $m+1$, including NOTA)
Each voter submits it's preferred candidate, the mechanism picks a winner.
Ideally no money changes hands.

Let's talk about ordinal mechanisms w/o money.

m alternatives, n voters, each voter i has a total order π_i over the alternatives
π_i is also called a "preference"
$\pi = (\pi_1, \pi_2, \dots, \pi_n)$ is a "preference profile"

Defn.: A Social Welfare Fn. $F: (\pi_1, \dots, \pi_n) \to \sigma$ (where σ is a total order over A)
A Social Choice Fn $f: (\pi_1, \dots, \pi_n) \to A$
i.e., an SWF outputs a ranking over alternatives
an SCF outputs a single alternative

Condorcet's Paradox

Consider an election w/ 3 alternatives a, b, c, 3 voters

$$\pi_1(a) > \pi_1(b) > \pi_1(c)$$
$$\pi_2(b) > \pi_2(c) > \pi_2(a)$$
$$\pi_3(c) > \pi_3(a) > \pi_3(b)$$

Any alternative chosen by an SCF will displease a majority of voters ...

Defn: Plurality
Given $\pi = (\pi_1, \dots, \pi_n)$,
say agent i "votes" for $a \in A$ if $\pi_i(a) > \pi_i(a'), \forall a' \neq a$
votes $(a) = |\{i : i \text{ votes for } a\}|$
The plurality is the SCF that chooses an alternative w/ maximum # votes.

We'll talk more about SCFs, for now let's talk about SWFs.

What are some "good" properties of SWFs?
Fix an SWF F. Let $\sigma = F(\pi_1, \dots, \pi_n)$

(i) Unanimity: if $\exists\, a, b \in A$ s.t.
$\forall i, \ \pi_i(a) > \pi_i(b)$, then $\sigma(a) > \sigma(b)$

π_1 : x x x a x x b x
π_2 : x a x x x b x x
π_3 : a x x x x x x b
σ : ... a ... b ...

(ii) (bad property) Dictatorship:
$\exists i \ \forall \pi, \ F(\pi) = \pi_i$

(iii) Independence of Irrelevant Alternatives

π	π'
π_1: .... a ... b	π_1': a .... b
π_2: .... b a	π_2': ... b ..... a .....
π_3: ... b .... a	π_3': .... b ... a .....
$F(\pi)$: .... b .... a .... $= \sigma$	$F(\pi')$: ... b .... a $= \sigma'$

Consider $\pi = (\pi_1, \dots, \pi_n)$, $\pi' = (\pi_1', \dots, \pi_n')$, $a, b \in A$
s.t. $\forall i$, $\pi_i(a) > \pi_i(b)$ iff $\pi_i'(a) > \pi_i'(b)$
then $\sigma(a) > \sigma(b)$ iff $\sigma'(a) > \sigma'(b)$
where $\sigma = F(\pi)$, $\sigma' = F(\pi')$

For $|A| \geq 3$,
Arrow's Impossibility Theorem (1950): Any SWF that satisfies Unanimity & IIA must be a dictatorship

What about if $|A| = 2$:
Consider the SWF that chooses the order preferred by at least $\lceil n/2 \rceil$ agents.
- easily seen to be unanimous
- if each agents orders a, b the same way in π & π', then $\pi = \pi'$, hence $F(\pi) = F(\pi')$

Suppose we extend to $|A| \geq 3$: choose the order preferred by a plurality of voters

π	π'
π_1, π_2, π_3: $a > b > c$	π_1': $a > b > c$
	π_2': $a > c > b$
	π_3': $c > a > b$
π_4, π_5: $b > c > a$	π_4', π_5': $b > c > a$
σ: $a > b > c$	σ': $b > c > a$

This does not satisfy IIA, since in π & π', each agent has the same order for a & b, however in σ & σ', a & b have different orders for a & b.

Step 1: IIA + Unanimity $\Rightarrow$ Pairwise Independence

Defn: Pairwise Neutrality
Given π, π', if $\exists\, a, b, c, d \in A$ s.t.
$\forall i \quad \pi_i(a) > \pi_i(b) \Leftrightarrow \pi_i'(c) > \pi_i'(d)$
then $\sigma(a) > \sigma(b) \Leftrightarrow \sigma'(c) > \sigma'(d)$

π	π'
π_1: a b c d e	π_1': b e a c d
π_2: b c d a e	π_2': e b a d c
π_3: c e a d b	π_3': c b e a d
σ: a b c d e	σ': ... c ... d ...

Q. what does IIA say about a & b? b & c (a, d) & (b, e)?

Claim: IIA + Unanimity $\Rightarrow$ PN

Proof: Given π, π' that satisfy conditions of the claim.
(we allow $\{a,b\} \cap \{c,d\} \neq \emptyset$), say $\sigma(a) > \sigma(b)$

π	π'	π''
π_1: .. a ... b ...	π_1': ... c ... d ...	π_1'': ... c a ... b d ...
π_2: ... b ... a ...	π_2': ... d ... c	π_2'': ... b d ... c a ...
π_3: ... b ... a ...	π_3': ... d ... c ...	π_3'': ... b d ... c a ...
π_4: ... a ... b ...	π_4': .... c ... d ...	π_4'': ... c a ... b d ...
σ: ... a ... b ...		

Construct π'':
$\forall i: \pi_i(a) > \pi_i(b) \Rightarrow \pi_i'(c) > \pi_i'(d)$
$\Rightarrow \pi_i''(c) > \pi_i''(a) > \pi_i''(b) > \pi_i''(d)$
& $\pi_i(b) > \pi_i(a) \Rightarrow \pi_i'(d) > \pi_i'(b)$
$\Rightarrow \pi_i''(b) > \pi_i''(d) > \pi_i''(c) > \pi_i''(a)$

Then $\sigma(a) > \sigma(b) \Rightarrow \sigma''(a) > \sigma''(b)$ by IIA
$\Rightarrow \sigma''(c) > \sigma''(d)$ by unanimity
$\Rightarrow \sigma'(c) > \sigma'(d)$ by IIA ∎

Step 2: Find a likely dictator.

Assume agents are indexed $1 \dots n$
Pick any $a, b \in A$. Let $\pi: \forall i, \ \pi_i(a) > \pi_i(b)$
$\pi': \forall i \ \pi_i'(b) > \pi_i'(b)$
Let $\pi^{(i)} = (\pi_1', \pi_2', \dots, \pi_i', \pi_{i+1}, \dots, \pi_n)$
So $\pi^{(0)} = \pi \quad (a > b \ \forall i)$
$\pi^{(n)} = \pi' \quad (b > a \ \forall i)$
In $\pi^{(i)}$, the first i voters prefer b to a, the remaining prefer a to b.

By unanimity, $F(\pi^{(0)})$: $a > b$
$F(\pi^{(n)})$: $b > a$
Hence $\exists k$: $F(\pi^{(k-1)})$: $a > b$
$F(\pi^{(k)})$: $b > a$
Voter k is our likely dictator

Step 3: Use pairwise neutrality to show that k must be a dictator

Fix any preference profile τ, and alternatives c, d
Let $\sigma = F(\tau)$
WTS: $\tau_k(c) > \tau_k(d) \Rightarrow \sigma(c) > \sigma(d)$
(and hence, k is a dictator)

Construct τ': pick some $e \neq c, d$

τ_1: ... c ... d ...	τ_1': e ... c ... d ...
τ_2: ... d ... c ...	τ_2': e ... d ... c ...
⋮	
τ_{k-1}: ... c ... d ...	τ_{k-1}': e ... c ... d ...
τ_k ... c ... d ...	τ_k': ... c ... e ... d ...
τ_{k+1} ... d ... c ...	τ_{k+1}': ... d ... c ... e
⋮	⋮
τ_n ... c ... d ...	τ_n': ... c ... d ... e

Clearly, $\sigma'(c) > \sigma'(d)$ iff $\sigma(c) > \sigma(d)$ by IIA

$\pi^{(k-1)}$	$\pi^{(k)}$
π_1': ... b ... a ...	π_1'
π_2': ... b ... a ...	π_2'
⋮	⋮
π_{k-1}': ... b ... a ...	π_{k-1}'
π_k: ... a ... b ...	π_k': ... b ... a ...
π_{k+1}: ... a ... b ...	π_{k+1}
⋮	⋮
π_n: ... a ... b ...	π_n
$\sigma^{(k-1)}$: ... a ... b ...	$\sigma^{(k)}$: ... b ... a ...

by PN: $\sigma'(c) > \sigma'(e)$
by PN: $\sigma'(e) > \sigma'(d)$
hence, $\sigma'(c) > \sigma'(d)$
$\Rightarrow \sigma(c) > \sigma(d)$ ∎