भारतीय प्रौद्योगिकी संस्थान मुंबई
Indian Institute of Technology Bombay

CS 6001: Game Theory and Algorithmic Mechanism Design

Week 7

Swaprava Nath

Slide preparation acknowledgments: C. R. Pradhit and Adit Akarsh

ज्ञानम् परमम् ध्येयम्
Knowledge is the supreme goal

Contents

Mechanism Design (Inverse Game Theory)

The objective/desired are set - the task is to set the rules of the game. Examples : Election, license scarce resource (spectrum, cloud), matching students to universities.

General Model

- N: set of players
- X: set of outcomes, e.g, winner in an election, which resource allocated to whom etc.
- Θ_i : set of private information of agent i **(type)**. A type $\theta_i \in \Theta_i$.
- The type may manifest in the preferences over the outcomes in different ways
  1. Ordinal : θ_i defines an ordering over the outcome.
  2. Cardinal : an utility function u_i maps an (outcome, type) pair to real numbers,
     - $u_i : X \times \Theta_i \to \mathbb{R}$ (private value model)
     - $u_i : X \times \Theta_i \to \mathbb{R}$ (interdependent value model)

Examples

Voting

- X is the set of candidates.
- θ_i is a ranking over this candidates, e.g., $\theta_i = (a, b, c)$, i.e., a is preferred more than b which is in turn more preferred than c.

Single Object allocation: an outcome is $x = (\underline{a}, \underline{p}) \in X$

- $\underline{a} = (a_1, a_2, \ldots, a_n)$, $a_i \in \{0, 1\}$, $\sum_{i \in N} a_i \leqslant 1$, allocations.
- $\underline{p} = (p_1, p_2, \ldots, p_n)$, p_i is the payment charged to i.
- θ_i : value of i for the object.
- $u_i(x, \theta_i) = a_i \theta_i - p_i$

Social Choice Function

- The designer has an objective and this is captured through a **Social Choice Function(SCF)**.

$$f : \Theta_1 \times \Theta_2 \times \ldots \times \Theta_n \to X$$

Examples

- in voting, if there is a candidate who beats everyone else in pairwise contests the he/she must be chosen as a winner.
- in public project choice, where $\theta_i : X \to \mathbb{R}$, value for each project pick, $f(\theta) \in \arg\max_{a \in X} \sum_{i \in N} \theta_i(a)$

Question

How can we create a game where $f(\theta)$ emerges as an outcome of an equilibrium?

Answer

We need mechanisms

Mechanisms

Definition

An indirect mechanism is a collection of message spaces and a decision rule $\langle M_1, M_2, \ldots, M_n, g\rangle$

- M_i is the message space of agent i
- $g : M_1 \times M_2 \times \ldots \times M_n \to X$

A direct mechanism is the same as above with $M_i = \Theta_i, \forall i \in N, g \equiv f$. The message space is similar to equipping every agent with a card deck and asking to pick some.

Question

Why these are not so commonplace?

Answer

Due to a result that will follow.

Weakly Dominant

Definition

In a mechanism $\langle M_1, M_2, \ldots, M_n, g \rangle$, a message m_i is **weakly dominant** for player i at θ_i if

$$u_i(g(m_i, \tilde{m}_{-i}), \theta_i) \geqslant u_i(g(m_i', \tilde{m}_{-i}), \theta_i), \forall \tilde{m}_{-i}, \forall m_i'$$

All subsequent definitions assume cardinal preferences, however they can be replaced with ordinal, e.g., the above one could be defined as

$$u_i(g(m_i, \tilde{m}_{-i}), \theta_i) \ \theta_i \ u_i(g(m_i', \tilde{m}_{-i}), \theta_i), \forall \tilde{m}_{-i}, \forall m_i'$$

Dominant Strategy Implementable (DSI)

Definition

An SCF $f : \Theta \to X$ is implemented in dominant strategies by $\langle M_1, M_2, \ldots, M_n, g\rangle$ if

- $\exists$ message mappings $s_i : \Theta_i \to M_i$, s.t, $s_i(\theta_i)$ is a dominant strategy for agent i at θ_i, $\forall \theta_i \in \Theta_i$, $\forall i \in \mathcal{N}$.
- $g(s_1(\theta_1), \ldots, s_n(\theta_n)) = f(\theta)$, $\forall \theta \in \Theta$

We call this an indirect implementation, i.e., SCF f is **dominant strategy implementable (DSI)** by $\langle M_1, M_2, \ldots, M_n, g\rangle$.

Dominant Strategy Incentive Compatible DSIC)

Definition

A direct mechanism $\langle \Theta_1, \Theta_2, \ldots, \Theta_n, f \rangle$ is **dominant strategy incentive compatible (DSIC)** if

$$u_i(g(\theta_i, \tilde{\theta}_{-i}), \theta_i) \geqslant u_i(g(\theta_i', \tilde{\theta}_{-i}), \theta_i), \forall \tilde{\theta}_{-i}, \theta_i', \theta_i, \forall i \in \mathcal{N}$$

To find if an SCF f is dominant strategy implementable, we need to search over all possible indirect mechanisms $\langle M_1, M_2, \ldots, M_n, g \rangle$. But luckily, there is a result that reduces the search space.

Contents

Relationship between DSI and DSIC

Revelation Principle (for DSI SCFs)

If there exists an indirect mechanism that implements f in dominant strategies, then f is DSIC.
Implication: Can focus on DSIC mechanisms WLOG.

Proof.

Let f is implemented by $\langle M_1, M_2, \ldots, M_n, g \rangle$, hence $\exists s_i : \Theta_i \to M_i$ s.t., $\forall i \in \mathcal{N}, \forall \tilde{m}_{-i}, m_i, \theta_i$,

$$u_i(g(s_i(\theta_i), \tilde{m}_{-i}), \theta_i) \geqslant u_i(g(m_i', \tilde{m}_{-i}), \theta_i) \quad (1)$$

$$g(s_i(\theta_i), s_{-i}(\theta_{-i}) = f(\theta_i, \theta_{-i}) \quad (2)$$

Eq. 1 holds for all $m_i', \tilde{m}_{-i}$, in particular, $m_i' = s_i(\theta_i'), \tilde{m}_{-i} = s_{-i}(\theta_{-i})$ where θ_i' and $\tilde{\theta}_{-i}$ are arbitrary. Hence,

$$u_i(g(s_i(\theta_i), s_{-i}(\theta_{-i})), \theta_i) \geqslant u_i(g(s_i(\theta_i'), s_{-i}(\theta_{-i})), \theta_i) \Rightarrow u_i(f(\theta_i, \tilde{\theta}_{-i}), \theta_i) \geqslant u_i(f(\theta_i', \tilde{\theta}_{-i}), \theta_i)$$

$\Rightarrow f$ is DSIC. □

Bayesian extension

- Agents may have probabilistic information about other's types.
- Types are generated from a common prior (common knowledge) and are revealed only to the respective agents.
- Recall : Bayesian games $\langle N, (M_i)_{i\in N}, (\Theta_i)_{i\in N}, P, (\Gamma_\theta)_{\theta\in\Theta}\rangle$

Bayesian extension

Definition

An (indirect) mechanism $\langle M_1, M_2, \dots, M_n, g \rangle$ implements an SCF f in a Bayesian equilibrium if

- $\exists$ a message mapping profile $(s_1, \dots, s_n)$, s.t., $s_i(\theta_i)$ maximizes the ex-interim utility of agent i, $\forall \theta_i$, $\forall i \in \mathbb{N}$, i.e.,

$$\mathbb{E}_{\theta_{-i}|\theta_i}[u_i(g(s_i(\theta_i), s_{-i}(\theta_{-i})), \theta_i)] \geqslant \mathbb{E}_{\theta_{-i}|\theta_i}[u_i(g(m_i', s_{-i}(\theta_{-i})), \theta_i)] \qquad \forall m_i', \forall \theta_i, \forall i \in \mathbb{N}$$

- $g(s_i(\theta_i), s_{-i}(\theta_{-i})) = f(\theta_i, \theta_{-i})$, $\forall \theta$

We call f is Bayesian implementable via $\langle M_1, M_2, \dots, M_n, g \rangle$ under the prior P.

Lemma

If an SCF f dominant strategy implementable, then it is Bayesian implementable.

Proof : Homework

Bayesian Incentive Compatible

Definition

A direct mechanism $\langle \Theta_1, \Theta_2, \ldots, \Theta_n, f \rangle$ is **Bayesian Incentive Compatible (BIC)** if $\forall \theta_i, \theta_i', \forall i \in \mathbb{N}$

$$\mathbb{E}_{\theta_{-i}|\theta_i}[u_i(f(\theta_i, \theta_{-i}), \theta_{-i}), \theta_i] \geqslant \mathbb{E}_{\theta_{-i}|\theta_i}[u_i(f(\theta_i', \theta_{-i}), \theta_{-i}), \theta_i]$$

Revelation Principle for BI SCFs

Revelation Principle (for BI SCFs)

If an SCF f is implementable in Bayesian equilibrium, then f is BIC.

- Proof idea is similar to the DSI, with expected utilities at appropriate places.
- For truthfulness of these two kinds, we will only consider incentive compatibility.
- These results hold even for ordinal preferences and mechanisms.

Contents

Arrow's Social Welfare Function Setup

Question

Ignoring the truthful revelation for a moment, can we reasonably aggregate opinions for a general setup?

Objective: create **social preferences** from **individual preferences**

- Finite set of alternatives $A = \{a_1, a_2, \ldots, a_m\}$
- Finite set of players $N = \{1, 2, \ldots, n\}$
- Each player i has a preference order R_i over A (A binary relation over A, aR_ib means alternative a is at least as good as b to i
- Properties of R_i
  1. **Completeness:** for every pair of alternatives $a, b \in A$, either aR_ib or bR_ia or both
  2. **Reflexivity:** $\forall a \in A$, aR_ia
  3. **Transitivity:** if aR_ib and bR_ic, then aR_ic, $\forall a, b, c \in A$ and $i \in N$

Arrow's Social Welfare Function Setup

- Set of all preference ordering is denoted by $\mathcal{R}$
- An ordering R_i is **linear** if for every $a, b \in A$ s.t. aR_ib and bR_ia implies $a = b$ (**Antisymmetric**)
- Set of all **linear** preference ordering is denoted by $\mathcal{P}$
- Any arbitrary ordering R_i can be decomposed into its
  (a) **asymmetric** part P_i
  (b) **symmetric** part I_i
- Example:

$$R_i = \begin{bmatrix} a \\ b, c \\ d \end{bmatrix} = \{(a,b), (a,c), (a,d), (b,c), (c,b), (b,d), (c,d)\}$$

$$\Rightarrow P_i = \begin{bmatrix} a & a \\ b & c \\ d & d \end{bmatrix} = \{(a,b), (a,c), (a,d), (b,d), (c,d)\}, \qquad I_i = \{(b,c), (c,b)\}$$

Arrovian Social Welfare Function (ASWF)

$$F : \mathcal{R}^n \rightarrow \mathcal{R} \qquad \text{domain and co-domain are both rankings}$$

- Motivation: the function F captures the collective ordering of the society, if the most preferred is not feasible, the society can move to the next and so on
- $F(R) = F(R_1, R_2, \ldots, R_n)$ is an ordering over the alternatives
- $\hat{F}(R)$ is the **asymmetric** part of $F(R)$
- $\overline{F}(R)$ is the **symmetric** part of $F(R)$

Pareto or Unanimity

Definition (Weak Pareto)

An ASWF F satisfies **weak Pareto** if $\forall a, b \in A$ and for every strict preference profile P, if aP_ib forall $i \in N$, then $a\hat{F}(R)b$.

Important: there can be P's where the 'if' condition does not hold, then the implication is **vacuously** true

Definition (Strong Pareto)

An ASWF F satisfies **strong Pareto** if $\forall a, b \in A$ and for every preference profile R, if aR_ib forall $i \in N$ and aP_jb for some $j \in N$, then $a\hat{F}(R)b$.

Question

Which property implies the other?

Independence of Irrelevant Alternatives

- We say $R_i, R'_i \in \mathcal{R}$ **agree** on $\{a, b\}$ for agent i if

$$aP_ib \Leftrightarrow aP'_ib,\; bP_ia \Leftrightarrow bP'_ia,\; aI_ib \Leftrightarrow aI'_ib$$

- We use the shorthand $R_i|_{a,b} = R'_i|_{a,b}$ to denote this for agent i
- If this holds for every $i \in N$, $R|_{a,b} = R'|_{a,b}$

Definition (Independence of Irrelevant Alternatives)

An ASWF F satisfies **independence of irrelevant alternatives** (IIA) if for all $a, b \in A$, and for every pair of preference profiles R and R', if $R|_{a,b} = R'|_{a,b}$, then $F(R)|_{a,b} = F(R')|_{a,b}$.

If the relative positions of two alternatives are the same in two different preference profiles, then the aggregate should also match the relative positions of those two alternatives

Example

If the relative positions of two alternatives are the same in two different preference profiles, then the aggregate should also match the relative positions of those two alternatives

R					R'			
a	a	c	d		d	c	b	b
b	c	b	c		a	a	c	a
c	b	a	b		b	b	a	d
d	d	d	a		c	d	d	c

- IIA says $F(R)|_{a,b} = F(R')|_{a,b}$
- Simple aggregation rules, e.g., **scoring rules**: each position of each agent gets a score $(s_1, s_2, \ldots, s_m), s_i \geqslant s_{i+1}, i = 1, 2, \ldots, m-1$, the final ordering is in the decreasing order of the scores
- One special scoring rule: **plurality**, $s_1 = 1, s_i = 0, i = 2, \ldots, m$.

Satisfaction of IIA

Question

Does plurality satisfy IIA?

R			
a	a	c	d
b	c	b	c
c	b	a	b
d	d	d	a

R'			
d	c	b	b
a	a	c	a
b	b	a	d
c	d	d	c

Check: $aF_{\texttt{plu}}(R)b$, but $bF_{\texttt{plu}}(R')a$, even though $R|_{a,b} = R'|_{a,b}$

Question

Does **dictatorship** satisfy IIA?

A **dictatorship** ASWF is where there exists a pre-determined agent d and $F^d(R) = R_d$

Arrow's impossibility result

Theorem (Arrow 1951)

For $|A| \geqslant 3$, if an ASWF F satisfies WP and IIA, then it must be dictatorial.

We cannot aggregate reasonably even when there is no truthfulness constraint

Contents

Decisiveness

Definition

Let $F : \mathcal{R}^n \to \mathcal{R}$ be given, $G \subseteq N, G \neq \emptyset$.

1. G is **almost decisive** over $\{a, b\}$ if for every R satisfying

$$aP_ib, \forall i \in G, \quad bP_ja, \forall j \in N \setminus G$$

   we have $a\hat{F}(R)b$.
   We will write this with the shorthand $\overline{D}_G(a, b)$: G is almost decisive over $\{a, b\}$ w.r.t. F.

2. G is **decisive** over $\{a, b\}$ if for every R satisfying

$$aP_ib, \forall i \in G$$

   we have $a\hat{F}(R)b$.
   We will write this with the shorthand $D_G(a, b)$: G is almost decisive over $\{a, b\}$ w.r.t. F.

Observation: $D_G(a, b) \Rightarrow \overline{D}_G(a, b)$

Proof of Arrow's theorem

The proof proceeds in two parts:

Part 1 **Field expansion lemma**: If a group is almost decisive over a pair of alternatives, it is decisive over all pairs of alternatives

Part 2 **Group contraction lemma**: If a group is decisive, then a strict non-empty subset of that group is also decisive.

Note: these two lemmas immediately proves the theorem

Field expansion lemma

Lemma

Let F satisfy WP and IIA, then $\forall a,b,x,y,G \subseteq N, G \neq \emptyset, a \neq b, x \neq y$

$$\overline{D}_G(a,b) \Rightarrow D_G(x,y).$$

It implies that under WP and IIA, the two notions of decisiveness are equivalent.

Cases to consider (ordered for the convenience of the proof):

1. $\overline{D}_G(a,b) \Rightarrow D_G(a,y), y \neq a,b$
2. $\overline{D}_G(a,b) \Rightarrow D_G(x,b), x \neq a,b$
3. $\overline{D}_G(a,b) \Rightarrow D_G(x,y), x,y \neq a,b$
4. $\overline{D}_G(a,b) \Rightarrow D_G(x,a), x \neq a,b$
5. $\overline{D}_G(a,b) \Rightarrow D_G(b,y), y \neq a,b$
6. $\overline{D}_G(a,b) \Rightarrow D_G(a,b)$
7. $\overline{D}_G(a,b) \Rightarrow D_G(b,a)$

Proof of FEL

- **Case 1:** $\overline{D}_G(a,b) \Rightarrow D_G(a,y), y \neq a, b$
- Pick an arbitrary $R \in \mathcal{R}^n$, s.t., aP_iy, $\forall i \in G$
- Need to show: $a\hat{F}(R)y$
- Construct R' s.t.

G		$N \setminus G$	
a	a	b	b
$\vdots$	$\vdots$	$\vdots$	$\vdots$
b	b	a	y
$\vdots$	$\vdots$	$\vdots$	$\vdots$
y	y	y	a

positions of a and y in $N \setminus G$ s.t. $R'|_{a,y} = R|_{a,y}$

- $\overline{D}_G(a,b) \Rightarrow a\hat{F}(R')b$
- WP over b, y, $\Rightarrow b\hat{F}(R')y$, transitivity $\Rightarrow a\hat{F}(R')y$
- IIA $\Rightarrow a\hat{F}(R)y$. Hence, $D_G(a,y)$

Proof of FEL (contd.)

- **Case 2:** $\overline{D}_G(a,b) \Rightarrow D_G(x,b), x \neq a,b$
- Pick an arbitrary $R \in \mathcal{R}^n$, s.t., xP_ib, $\forall i \in G$
- Need to show: $x\hat{F}(R)b$
- Construct R' s.t.

G		$N \setminus G$	
x	x	x	b
$\vdots$	$\vdots$	$\vdots$	$\vdots$
a	a	b	x
$\vdots$	$\vdots$	$\vdots$	$\vdots$
b	b	a	a

positions of x and b in $N \setminus G$ s.t. $R'|_{x,b} = R|_{x,b}$

- $\overline{D}_G(a,b) \Rightarrow a\hat{F}(R')b$
- WP over x, a, $\Rightarrow x\hat{F}(R')a$, transitivity $\Rightarrow x\hat{F}(R')b$
- IIA $\Rightarrow x\hat{F}(R)b$. Hence, $D_G(x,b)$

Proof of FEL (other cases)

- **Case 3:** $\overline{D}_G(a,b) \overset{\text{(case 1)}}{\Longrightarrow} D_G(a,y)\ (y \neq a,b) \overset{\text{(definition)}}{\Longrightarrow} \overline{D}_G(a,y) \overset{\text{(case 2)}}{\Longrightarrow} D_G(x,y)\ (x \neq a,y)$
- **Case 4:** $\overline{D}_G(a,b) \overset{\text{(case 2)}}{\Longrightarrow} D_G(x,b)\ (x \neq a,b) \overset{\text{(definition)}}{\Longrightarrow} \overline{D}_G(x,b) \overset{\text{(case 1)}}{\Longrightarrow} D_G(x,a)\ (x \neq a,b)$
- **Case 5:** $\overline{D}_G(a,b) \overset{\text{(case 1)}}{\Longrightarrow} D_G(a,y)\ (y \neq a,b) \overset{\text{(definition)}}{\Longrightarrow} \overline{D}_G(a,y) \overset{\text{(case 2)}}{\Longrightarrow} D_G(b,y)\ (y \neq a,b)$
- **Case 6:** $\overline{D}_G(a,b) \overset{\text{(case 2)}}{\Longrightarrow} D_G(x,b)\ (x \neq a,b) \overset{\text{(definition)}}{\Longrightarrow} \overline{D}_G(x,b) \overset{\text{(case 2)}}{\Longrightarrow} D_G(a,b)$
- **Case 7:** $\overline{D}_G(a,b) \overset{\text{(case 5)}}{\Longrightarrow} D_G(b,y)\ (y \neq a,b) \overset{\text{(definition)}}{\Longrightarrow} \overline{D}_G(b,y) \overset{\text{(case 1)}}{\Longrightarrow} D_G(b,a)$

Group contraction lemma

Lemma

Let F satisfy WP and IIA, and let $G \subseteq N, G \neq \emptyset, |G| \geqslant 2$ be decisive. Then $\exists G' \subset G, G' \neq \emptyset$ which is also decisive.

Proof:

- $G, |G| \geqslant 2$ is given. Let $G_1 \subset G, G_2 = G \setminus G_1, G_1, G_2 \neq \emptyset$, arbitrary.
- Construct R

G_1	G_2	$N \setminus G$
a	c	b
b	a	c
c	b	a

$$aP_i b, \forall i \in G \text{ and } G \text{ decisive } \Rightarrow a\hat{F}(R)b$$

- Where can c stand in $F(R)$ w.r.t. a? We will show in every possible case, either G_1 or G_2 will be decisive

Proof of GCL

Case 1: $a\hat{F}(R)c$

G_1	G_2	$N \setminus G$
a	c	b
b	a	c
c	b	a

have seen $\Rightarrow a\hat{F}(R)b$

- Consider G_1
- $aP_ic, \forall i \in G_1, cP_ia, \forall i \in N \setminus G_1$
- Consider each R' where the above relation holds
- by IIA $a\hat{F}(R')c$
- Hence $\overline{D}_{G_1}(a,c) \stackrel{\text{(FEL)}}{\Longrightarrow} D_{G_1}$

Proof of GCL (contd.)

Case 2: $\neg(a\hat{F}(R)c) \implies cF(R)a$

G_1	G_2	$N \setminus G$
a	c	b
b	a	c
c	b	a

have seen $\Rightarrow a\hat{F}(R)b$

- $a\hat{F}(R)b$ and $cF(R)a$ give $c\hat{F}(R)b$
- Consider G_2
- $cP_ib, \forall i \in G_2, bP_ic, \forall i \in N \setminus G_2$
- Consider each R' where the above relation holds
- by IIA $c\hat{F}(R')b$
- Hence $\overline{D}_{G_2}(c,b) \overset{\text{(FEL)}}{\Longrightarrow} D_{G_2}$

भारतीय प्रौद्योगिकी संस्थान मुंबई

Indian Institute of Technology Bombay