CS 886: Multiagent Systems

Introduction to Mechanism Design

Kate Larson

Computer Science
University of Waterloo

Outline

Gibbard-Satterthwaite Impossibility

Theorem

Assume that

- O *is finite and* $|O| \geq 3$,
- *each* $o \in O$ *can be achieved by SCF* f *for some* θ, *and*
- Θ *includes all possible strict orderings over* O.

Then f *is implementable in dominant strategies (strategy-proof) if and only if it is dictatorial.*

Definition

SCF f *is* **dictatorial** *if there is an agent* i *such that for all* θ

$$f(\theta) \in \{o \in O | u_i(o, \theta_i) \geq u_i(o', \theta_i) \forall o' \in O\}$$

Gibbard-Satterthwaite Impossibility

Theorem

Assume that

- *O is finite and $|O| \geq 3$,*
- *each $o \in O$ can be achieved by SCF f for some θ, and*
- *Θ includes all possible strict orderings over O.*

Then f is implementable in dominant strategies (strategy-proof) if and only if it is dictatorial.

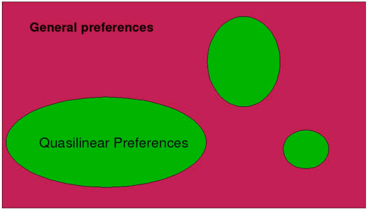
Definition

SCF f is **dictatorial** *if there is an agent i such that for all θ*

$$f(\theta) \in \{o \in O | u_i(o, \theta_i) \geq u_i(o', \theta_i) \forall o' \in O\}$$

Circumventing Gibbard-Satterthwaite

- Use a weaker equilibrium concept
- Design mechanisms where computing a beneficial manipulation is hard
- Randomization
- Restrict the structure of agents' preferences



1

Single-Peaked Preferences

Definition

A binary relation $\geq$ on a set of alternatives O is a linear order on O if it is reflexive, transitive, and total.

Definition

A preference relation $\succeq$ is single-peaked with respect to the linear order $\geq$ on O if there is an alternative $x \in O$ with the property that $\succeq$ is increasing with respect to $\geq$ on $\{y \in O | x \geq y\}$ and decreasing with respect to $\geq$ on $\{y \in O | y \geq x\}$. That is

$$\text{If } x \geq z > y \text{ then } z \succ y$$

and

$$\text{If } y > z \geq x \text{ then } z \succ y.$$

Single-Peaked Preferences

Definition

Let x_i denote agent $i \in A$'s "peak". Agent $h \in A$ is a median agent if

$$|\{a_i \in A | x_i \geq x_h\}| \geq \frac{|A|}{2} \text{ and } |\{a_i \in A | x_h \geq x_i\}| \geq \frac{|A|}{2}.$$

Quasi-linear preferences

- Outcome $o = (x, t_1, \dots, t_n)$
  - x is a "project choice"
  - $t_i \in \mathbb{R}$ are transfers (money)
- Utility function of agent i

$$u_i(o, \theta_i) = v_i(x, \theta_i) - t_i$$

- Quasi-linear mechanism

$$M = (S_1, \dots, S_n, g(\cdot))$$

where

$$g(\cdot) = (x(\cdot), t_1(\cdot), \dots, t_n(\cdot))$$

Quasi-linear preferences

- Outcome $o = (x, t_1, \ldots, t_n)$
  - x is a "project choice"
  - $t_i \in \mathbb{R}$ are transfers (money)
- Utility function of agent i

$$u_i(o, \theta_i) = v_i(x, \theta_i) - t_i$$

- Quasi-linear mechanism

$$M = (S_1, \ldots, S_n, g(\cdot))$$

where

$$g(\cdot) = (x(\cdot), t_1(\cdot), \ldots, t_n(\cdot))$$

Quasi-linear preferences

- Outcome $o = (x, t_1, \ldots, t_n)$
  - x is a “project choice”
  - $t_i \in \mathbb{R}$ are transfers (money)
- Utility function of agent i

$$u_i(o, \theta_i) = v_i(x, \theta_i) - t_i$$

- Quasi-linear mechanism

$$M = (S_1, \ldots, S_n, g(\cdot))$$

where

$$g(\cdot) = (x(\cdot), t_1(\cdot), \ldots, t_n(\cdot))$$

Social Choice Functions and Quasi-linearity

- SCF is **efficient** if for all θ

$$\sum_{i=1}^{n} v_i(x(\theta), \theta_i) \geq \sum_{i=1}^{n} v_i(x'(\theta), \theta_i) \forall x'(\theta)$$

  This is also known as **social welfare maximizing**

- SCF is **budget-balanced** if

$$\sum_{i=1}^{n} t_i(\theta) = 0$$

  Weakly budget-balanced if

$$\sum_{i=1}^{n} t_i(\theta) \geq 0$$

Social Choice Functions and Quasi-linearity

- SCF is **efficient** if for all θ

$$\sum_{i=1}^{n} v_i(x(\theta), \theta_i) \geq \sum_{i=1}^{n} v_i(x'(\theta), \theta_i) \forall x'(\theta)$$

  This is also known as **social welfare maximizing**

- SCF is **budget-balanced** if

$$\sum_{i=1}^{n} t_i(\theta) = 0$$

  Weakly budget-balanced if

$$\sum_{i=1}^{n} t_i(\theta) \geq 0$$

Groves Mechanisms [Groves 73]

A **Groves mechanism** $M = (S_1, \ldots, S_n, (x, t_1, \ldots, t_n))$ is defined by

- Choice rule

$$x^*(\theta) = \arg\max_x \sum_i v_i(x, \theta_i)$$

- Transfer rules

$$t_i(\theta) = h_i(\theta_{-i}) - \sum_{j \neq i} v_j(x^*(\theta), \theta_j)$$

where $h_i(\cdot)$ is an (arbitrary) function that does not depend on the reported type θ_i' of agent i.

Groves Mechanisms [Groves 73]

A **Groves mechanism** $M = (S_1, \ldots, S_n, (x, t_1, \ldots, t_n))$ is defined by

- Choice rule

$$x^*(\theta) = \arg\max_x \sum_i v_i(x, \theta_i)$$

- Transfer rules

$$t_i(\theta) = h_i(\theta_{-i}) - \sum_{j \neq i} v_j(x^*(\theta), \theta_j)$$

where $h_i(\cdot)$ is an (arbitrary) function that does not depend on the reported type θ_i' of agent i.

Groves Mechanisms [Groves 73]

A **Groves mechanism** $M = (S_1, \ldots, S_n, (x, t_1, \ldots, t_n))$ is defined by

- Choice rule

$$x^*(\theta) = \arg\max_x \sum_i v_i(x, \theta_i)$$

- Transfer rules

$$t_i(\theta) = h_i(\theta_{-i}) - \sum_{j \neq i} v_j(x^*(\theta), \theta_j)$$

where $h_i(\cdot)$ is an (arbitrary) function that does not depend on the reported type θ_i' of agent i.

Groves Mechanisms

Theorem

Groves mechanisms are strategy-proof and efficient.

We have gotten around Gibbard-Satterthwaite.

Proof

Agent i's utility for strategy $\hat{\theta}_i$, given $\hat{\theta}_{-i}$ from agents $j \neq i$ is

$$\begin{aligned} u_i(\hat{\theta}_i) &= v_i(x^*(\hat{\theta}, \theta_i)) - t_i(\hat{\theta}) \\ &= v_i(x^*(\hat{\theta}, \theta_i)) + \sum_{j \neq i} v_j(x^*(\hat{\theta}, \hat{\theta}_j) - h_i(\hat{\theta}_{-i}) \end{aligned}$$

Ignore $h_i(\hat{\theta}_{-i})$ and notice $x^*(\hat{\theta}) = \arg\max_x \sum_i v_i(x, \hat{\theta}_i)$ i.e it maximizes the sum of reported values. Therefore, agent i should announce $\hat{\theta}_i = \theta_i$ to maximize its own payoff.

Thm: Groves mechanisms are unique (up to $h_i(\theta_{-i})$).

Vickrey-Clarke-Groves Mechanism

aka Clarke mechansism, aka Pivotal mechanism

- Implement efficient outcome

$$x^* = \arg\max_x \sum_i v_i(x, \theta_i)$$

- Compute transfers

$$t_i(\theta) = \sum_{j \neq i} v_j(x^{-i}, \theta_j) - \sum_{j \neq i} v_j(x^*, \theta_j)$$

where $x^{-i} = \arg\max_x \sum_{j \neq i} v_j(x, \theta_j)$

VCG are efficient and strategy-proof.

Vickrey-Clarke-Groves Mechanism

aka Clarke mechansism, aka Pivotal mechanism

- Implement efficient outcome

$$x^* = \arg\max_x \sum_i v_i(x, \theta_i)$$

- Compute transfers

$$t_i(\theta) = \sum_{j \neq i} v_j(x^{-i}, \theta_j) - \sum_{j \neq i} v_j(x^*, \theta_j)$$

where $x^{-i} = \arg\max_x \sum_{j \neq i} v_j(x, \theta_j)$

VCG are efficient and strategy-proof.

VCG Mechanism

Agent's equilibrium utility is

$$
\begin{aligned}
u_i((x^*, t), \theta_i) &= v_i(x^*, \theta_i) - \left[\sum_{j \neq i} v_j(x^{-i}, \theta_j) - \sum_{j \neq i} v_j(x^*, \theta_j)\right] \\
&= \sum_{j=1}^{n} v_j(x^*, \theta_j) - \sum_{j \neq i} v_j(x^{-i}, \theta_j) \\
&= \text{marginal contribution to the welfare of the system}
\end{aligned}
$$

Examples

- Single item auction
- Public Good
- Multi-item auction

Characterization of Incentive Compatible Mechanisms

A mechanism is incentive-compatible if and only if it satisfies the following conditions for for every i and every v_{-i}

- The transfer, t_i does not depend on v_i, but only on the alternative chosen, $x(v_i, v_{-i})$.
- The mechanism optimizes for each agent. That is, for every v_i we have that $x(v_i, v_{-i}) \in \arg\max(v_i(x) - t_i)$.

Charcterization of Incentive Compatible SCF

Definition

A social choice function f satisfies weak monotonicity (WMON) if for all i and for all v_{-i} we have that $f(v_i, v_{-i}) = a \neq b = f(v_i', v_{-i})$ implies that $v_i(a) - v_i(b) \geq v_i'(a) - v_i'(b)$.

Theorem

If a mechanism is incentive compatible, then the social choice function it implements is WMON. If the domain of all agents' value functions is convex then for every social choice function that satisfies WMON, there exists transfers such that the resulting mechanism is incentive compatible.

Charcterization of Incentive Compatible SCF

Definition

A social choice function f satisfies weak monotonicity (WMON) if for all i and for all v_{-i} we have that $f(v_i, v_{-i}) = a \neq b = f(v_i', v_{-i})$ implies that $v_i(a) - v_i(b) \geq v_i'(a) - v_i'(b)$.

Theorem

If a mechanism is incentive compatible, then the social choice function it implements is WMON. If the domain of all agents' value functions is convex then for every social choice function that satisfies WMON, there exists transfers such that the resulting mechanism is incentive compatible.

Affine Maximizers

A social choice function f is an *affine maximizer* if given agent weights $a_i \in \mathbb{R}^+$ and outcome weights w_o, $o \in O$

$$f(v_1, \dots, v_n) \in \arg\max(w_o + \sum_i a_i v_i(o)).$$

Theorem (Roberts)

If $|O| \geq 3$ and f is onto and we place no restrictions on v_i for all i, and $(f, t_1, \dots, t_n)$ is incentive compatible, then f is an affine maximizer.

Affine Maximizers

A social choice function f is an *affine maximizer* if given agent weights $a_i \in \mathbb{R}^+$ and outcome weights w_o, $o \in O$

$$f(v_1, \ldots, v_n) \in \arg\max(w_o + \sum_i a_i v_i(o)).$$

Theorem (Roberts)

If $|O| \geq 3$ and f is onto and we place no restrictions on v_i for all i, and $(f, t_1, \ldots, t_n)$ is incentive compatible, then f is an affine maximizer.

Single-Parameter Domains

An agent has a *single-parameter* preference if it has one value if it "wins" and zero if it "loses". That is, all "winning" alternatives are equivalent.

- **Monotonicity:** Let W_i be the set of outcomes that i considers to be "winning". A SCF F is monotone in v_i if $\forall v_{-i}$, and every $v_i \leq v_i'$ we have that $f(v_i, v_{-i}) \in W_i$ implies that $f(v_i', v_{-i}) \in W_i$.
- **Critical Value:** $c_i(v_{-i}) = \sup_{v_i | f(v_i, v_{-i}) \notin W_i} v_i$.

Mechanisms for Single-Parameter Domains

A mechanism $(f, t_1, \ldots, t_n)$ on a single parameter domain is incentive compatible if and only if the following hold:

- f is monotone in every v_i
- Every winning bid pays its critical value, and losing bids pay zero.