MECHANISM DESIGN

Mechanism Design

- "The mechanism design problem is to implement an optimal system-wide solution to a decenralized optimization problem with self-interested agents with private information about their preferences for different outcomes".
- Two ways to think of it:
  - How to get rugged individualists to work together.
  - How to design the payoff matrix so that agents pick the strategies that you want them to pick.
- Often described as setting the rules of the game.
- This material is taken from David Parkes' thesis "Iterative Combinatorial Auctions", U. Penn, 2001.

Types

- We met types at the end of the last class.
- Captures the uncertainty that agents have about the payoffs of other agents.
- $\theta_i \in \Theta_i$ denotes the type of i.
- For every outcome $o \in \mathcal{O}$ i has a utility dependent on its type $u_i(o, \theta_i)$.
- i prefers o_1 to o_2 when

$$u_i(o_1, \theta_i) > u_i(o_2, \theta_i)$$

Utility

- Usual assumption is that utilities are *quasi-linear*.

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

  where o defines a choice $x \in \mathcal{K}$ from some discrete choice set and a payment p_i.

- In the kind of allocation problems we deal with in auctions, the $\mathcal{K}$ are allocations of goods, and the payments are to the auctioneer.
- Quasi-linear preferences make it easy to distribute utility between agents.

Strategies

- A strategy gives the action(s) an agent will adopt in every possible state of the world.
- $s_i(\theta_i) \in \Sigma_i$ denotes the strategy i adopts, out of the set of all possible strategies Σ_i when it has type θ_i.
- In an English auction setting,
  - World state is (p, X), where p is current price, $X = (x_1, \ldots, x_n)$, and $x_i \in \{0, 1\}$ denotes whether or not agent i is the high bidder on the item.
  - A strategy defines a bid $b(p, X, v_i)$ for every (p, X) and every value v_i that i might hold.
  - A best response strategy for i is:

$$b_{BR}(p, X, v_i) = \begin{cases} p & \text{if } x_i = 0 \text{ and } p < v_i \\ \text{no bid} & \text{otherwise} \end{cases}$$

Social choice

- The notion of a *social choice function* defines the system-wide goal in mechanism design.
- A social choice function:

$$f : \Theta_1 \times \ldots \times \Theta_n \mapsto \mathcal{O}$$

chooses an outcome $f(\theta) \in \mathcal{O}$ given a set of types $\theta = (\theta_1, \ldots, \theta_n)$.

- Thus, given the types θ, we want to be able to choose the outcome, and the social choice function captures the relationship between the two.

Mechanism

- Formally a *mechanism*

$$\mathcal{M} = (\Sigma_1, \ldots, \Sigma_n, g(\cdot))$$

defines the set of strategies Σ_i available to each agent, and an *outcome rule*:

$$g : \Sigma_1 \times \ldots \times \Sigma_n \mapsto \mathcal{O}$$

such that $g(s)$ is the outcome implemented by the mechanism for strategy profile:

$$s = (s_1, \ldots, s_n)$$

- Thus the mechanism defines the available strategies and the rule for determining the outcome based on the strategies that the agents choose.

So

- We have a mechanism, which defines the strategies and the way that the outcome is computed.
  - This is the bit we can control.
  - "You are allowed to stop the clock at any time. First person to stop the clock wins, and pays the price on the clock".
- We have a social choice function, which defines the relationship between types and outcome.
  - This specifies what we want to happen.
  - "The winner should be the bidder with the highest valuation for the good"
- The idea that connects them is *implementation*.

Implementation

- A mechanism
$$\mathcal{M} = (\Sigma_1, \ldots, \Sigma_n, g(\cdot))$$
implements a social choice function $f(\theta)$ if:
$$f(\theta) = g(s_1^*(\theta_1), \ldots, s_n^*(\theta_n))$$
for all
$$(\theta_1, \ldots, \theta_n) \in \Theta_1 \times \ldots \times \Theta_n$$
where strategy profile $(s_1^*, \ldots, s_n^*)$ is an equilibrium solution to the game induced by $\mathcal{M}$.
- Thus a mechanism implements a social choice function if the outcome of the mechanism in equilibrium is that specified by the social choice function for all possible agent preferences.

Effect of utilities

- Since utilities are quasi-linear, we can reduce a social choice function to

$$f(\theta) = (x(\theta), p_1(\theta), \ldots, p_n(\theta))$$

  a choice $x(\theta) \in \mathcal{K}$ and a payment made by each agent.

- We can then decompose the outcome rule $g(s)$ for a mechanism into:
  - a *choice rule*, $k(s)$, which makes a choice based on the set of strategies the agents pick, and
  - a *payment rule*, $t_i(s)$, which tells each agent what it is going to have to pay.

Properties of social choice mechanisms

- A social choice function is Pareto optimal if for every $o' \in f(\theta)$ and for all θ:

$$u_i(o', \theta_i) > u_i(o, \theta_i) \Rightarrow \exists j, u_j(o', \theta_j) < u_j(o, \theta_j)$$

- We can also exploit the separability of the choice and payment parts of a quasi-linear mechanism to think of the properties of social choice functions in those terms, seperately.
- A social choice function is *allocatively efficient* if, for all θ:

$$\sum_i v_i(x(\theta), \theta_i) \geq \sum_i v_i(x', \theta_i)$$

  for all $x' \in \mathcal{K}$.

- An efficient allocation maximises the total value over all agents.

- A social choice function is *budget balanced* if, for all θ

$$\sum_i p_i(\theta) = 0$$

in other words if there are no payments into or out of the mechanism.

- A mechanism that is allocatively efficient and budget-balanced is Pareto optimal.
- A social choice function is *weak budget balanced* if, for all θ

$$\sum_i p_i(\theta) \geq 0$$

in other words there can be payments from the agents to the mechanism, but not from the mechanism to the agents.

Properties of mechanisms

- A mechanism is Pareto optimal, if it implements a Pareto optimal social choice function $f(\theta)$.
- This is *ex post* Pareto optimality — it works for a specific set of agent types.
- There is a (weaker) *ex ante* notion where the relationship holds in expectation of the types.
  - No outcome that one agent strictly prefers and all others weakly prefer.
- A mechanism is efficient if it implements an allocatively efficient social choice function $f(\theta)$.
- We can also say things about budget-balance.

- A mechanism is *ex ante* budget-balanced if the equilibrium net transfers to the mechanism are balanced *in expectation* for a distribution over agent preferences.
- A mechanism is *ex post* budget-balanced if the equilibrium net transfers to the mechanism are non-negative for *all* agent preferences.

- These measures all talk about the mechanism from the outside.
- We can also think about it from the perspective of an individual agent.
- A mechanism is *interim* individual-rational if, for all preferences θ, it implements a social choice mechanism $f(\theta)$ with:

$$u_i(f(\theta_i, \theta_{-i}) \geq \overline{u}_i(\theta_i))$$

  where $u_i(f(\theta_i, \theta_{-i})$ is the expected utility for i at the outcome based on the distribution of θ_{-i}, and $\overline{u}_i(\theta_i)$ is the expected utility for non-participation.
- In other words, a mechanism is individually rational if an agent can expect to do better by engaging in it than not.
- The "interim" bit acknowledges the agent knows its type, but just has knowledge of the distribution of other agents' types.

- Also *ex post* IR, where expected utility of participating is at least as good as non-participation for all possible types of other agents.
- Also *ex ante* IR where expected utility of participating, averaged over all its possible types and all other agents' possible types, is at least as good as non-participation.

Revelation principle

- A *direct-revelation* mechanism restricts the strategy set $\Sigma_i \in \Theta_i$ for all i to be the reporting of the agent's type and selects an outcome $g(\hat{\theta})$ based on the reported types $\hat{\theta} = (\hat{\theta}_i, \ldots, \hat{\theta}_n)$
- In other words each i reports a type $\hat{\theta}_i = s_i(\theta_i)$ based on its type θ_i.
- A strategy is *truth-revealing* if it reports true information about preferences.
- A mechanism is *incentive compatible* if the equilibrium strategy profile has every agent reporting its true preferences.

- Different flavors of incentive compatibility.
- A mechanism is Bayesian-Nash incentive-compatible if truth-revelation is a Bayesian-Nash equilibrium of the game induced by the mechanism.
- That is in such a mechanism the strategy that maximises every agent's expected utility is to truthfully report its preferences provided all other agents do the same.
- There is an even stronger notion of incentive-compatibility.
- A mechanism is *strategy proof* if truth revelation is a dominant strategy equilibrium.
- This kind of mechanism is very desirable.

- In an incentive-compatible mechanism, the outcome rule is exactly the social choice rule that the mechanism implements.
- An incentive-compatible direct-revelation mechanism implements the social choice function $f(\theta) = g(\theta)$ where $g(\theta)$ is the outcome rule of the mechanism.
- Groves mechanisms are such mechanisms, and they compute efficient allocations under conditions that make truth-revelation the dominant strategy.

- Okay, now the revelation principle itself.
- Suppose there exists a mechanism, direct or otherwise, that implements the social choice function $f(\cdot)$ in dominant strategies. Then $f(\cdot)$ is truthfully implementable in dominant strategies.
- In other words there is a strategy-proof mechanism that implements $f(\cdot)$.
- Suppose there exists a mechanism, direct or otherwise, that implements the social choice function $f(\cdot)$ in Bayesian-Nash equilibrium. Then $f(\cdot)$ is truthfully implementable in a Bayeisan-Nash incentive compatible direct-revelation mechanism.
- The distribution of agent types has to be common knowledge to the mechanism as well as the agents.

Vickrey/Clarke/Groves

- Vickrey-Clarke-Groves mechanisms for quasilinear preferences are mechanisms that are allocatively efficient, strategy-proof and direct revelation.
  - Also called VCG, "Groves mechanisms"
- Some special cases are also weakly budget-balanced, so there is no need for an external subsidy.
- These mechanisms are the *only* ones for quasi-linear preferences and general valuation functions that are allocatvely efficient, strategy-proof and direct revelation.

- In a Groves mechanism i reports its type $\hat{\theta}_i = s_i(\theta_i)$ which may not be its true type.
- Given the reported types $\hat{\theta} = (\theta_1, \ldots, (\theta_n)$ the VCG choice rule computes:

$$k^*(\hat{\theta}) = \arg\max_{k \in \mathcal{K}} \Sigma_i v_i(k, \hat{\theta}_i)$$

  This is the choice that maximises the total reported value over all agents.
- The payment rule is then:

$$t_i(\hat{\theta}) = h_i(\hat{\theta}_{-i}) - \Sigma_{j \neq i} v_j(k^*, \hat{\theta}_j)$$

  where $h_i : \Theta_{-i} \mapsto \Re$ is an arbitrary function on the reported types — this allows for a family of mechanisms with different tradeoffs across budget-balance and individual rationality.

Examples

- Vickrey auction
- Generalized Vickrey auction
  - Combinatorial allocation
- Pivotal/Clarke mechanism
  - Payment set to ensure individual rationality.

Impossibility results

- Gibbard-Satterthwaite.
- If agents have general preferences, and there are at least two agents, and at least three different optimal outcomes over the set of all agent preferences, then a social choice function is dominant-strategy implementable if and only if it is dictatorial.
- Escape these bounds by being more specific.
- Thus in markets, with quasi-linear preferences, we are not bound by it.

- There are problems even with quasi-linear preferences.
- A *simple exchange* is one with buyers and sellers selling single units of the same good.
- Hurwicz
- It is impossible to implement an efficient budget-balanced and strategy-proof mechanism in a simple exchange exconomy with quasi-linear preferences.
- This is bad news for auctions.
- (Equally you can show that for certain kinds of simple exchange, the expected profit for not truth-telling rapidly declines to zero as the number of buyers and sellers increases.)

- Can extend Hurwicz from strategy-proof, which implies implementable in dominant strategies, to Bayesian-Nash implementation.
- Myerson-Satterthwaite.
- It is impossible to achieve allocative efficiency, budget-balance and interim individual rationality in a Bayesian-Nash incentive-compatible mechanism, even with quasi-linear preferences.
- So we can hope to get at most two of:
  - Efficiency
  - Individual rationality; and
  - Budget balance

Summary

- This was a brief introduction to mechanism design.
- Covered the basic ideas.
- Talked about the revelation principle.
- Discussed VCG.
- Finished up with the impossibility results.