6.254 : Game Theory with Engineering Applications
Lecture 20: Mechanism Design II

Asu Ozdaglar
MIT

May 4, 2010

Outline

- Mechanism design from social choice point of view
- Implementation in dominant strategies
- Revelation principle
- VCG Mechanisms and examples
- Budget-balancedness
- dAGV Mechanisms

- Reading:
- Microeconomic Theory, MasColell, Whinston and Green, Chapter 23.
- Algorithmic Game Theory, edited by Nisan, Roughgarden, Tardos, and Vazirani, Chapter 9, by Noam Nisan.

Introduction

- Our goal is to analyze how individual preferences can be aggregated into desirable social or collective decisions.
- An important feature of such settings in which collective decisions must be made is that individuals' actual preferences are not publicly observable.
- As a result, in one way or another, individuals must be relied upon to reveal this information $\Rightarrow$ mechanism design problem.

Model

- Consider a setting with I agents.
- These agents must make a collective choice from some set Y (of possible alternatives).
- Each agent privately observes his preferences over the alternatives in Y.
- We model this by assuming that agent i privately observes a signal θ_i that determines his preferences, i.e., θ_i is agent i's type. The set of possible types for agent i is denoted by Θ_i and we use the notation $\Theta = \prod_{i=1}^{I} \Theta_i$.
- Each agent i is assumed to be an expected utility maximizer, whose Bernoulli utility function is given by $u_i(y, \theta_i)$ with $y \in Y$.
- As in all incomplete information settings, we assume that agents' types are drawn from a commonly known prior distribution over the type $\theta = (\theta_1, \ldots, \theta_I)$, and the type distribution and the utility functions $u_i(y, \theta_i)$ are common knowledge among the agents.

Social Choice Function

Since agents' preferences depend on the realization of their types, the agents may want the collective decision to depend on θ.

Definition

A social choice function is a function $f : \Theta_1 \times \cdots \times \Theta_I \to Y$ *that for each possible profile of agents' types* $(\theta_1, \ldots, \theta_I)$, *assigns a collective choice* $f(\theta_1, \ldots, \theta_I) \in Y$.

Definition

The social choice function f *is ex-post efficient if for no profile* $\theta = (\theta_1, \ldots, \theta_I)$ *is there a* $y \in Y$ *such that* $u_i(y, \theta_i) \geq u_i(f(\theta), \theta_i)$ *for all* i, *and* $u_i(y, \theta_i) > u_i(f(\theta), \theta_i)$ *for some* i.

- The problem is that the θ_i's are not publicly observable, so for the social choice $f(\theta_1, \ldots, \theta_I)$ to be chosen, each agent must be relied on to disclose their type correctly. But for a given $f(\cdot)$ an agent may not find it in his best interest to reveal this information truthfully.

Mechanism

Definition

A mechanism $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ is a collection of I strategy sets $(S_1, \ldots, S_I)$ and an outcome function $g : S_1 \times \cdots \times S_I \rightarrow Y$.

- A mechanism can be viewed as an institution with rules governing the procedure for making the collective choice:
  - The allowed actions are given by the strategy set S_i and the rule for how agents' actions get turned into a social choice is given by the outcome function g.
- The mechanism Γ combined with possible types $(\Theta_1, \ldots, \Theta_I)$, type distribution, and the utility functions u_i defines a Bayesian game, where
  - Payoffs are given by $u_i(g(s_1, \ldots, s_I), \theta_i)$,
  - A strategy is a function $s_i : \Theta_i \rightarrow S_i$.
- We say that a mechanism implements social choice function $f(\cdot)$ if there is an "equilibrium" of the game induced by the mechanism that yields the same outcomes as $f(\cdot)$ for each possible profile of types θ.

Dominant Strategy Implementation

- The mechanism design literature has investigated the implementation question for a variety of solution concepts.
- We focus on two solution concepts: dominant strategy equilibrium and Bayesian Nash equilibrium.

Definition

We say that a strategy profile $s^*(\cdot)$ is a (weakly) dominant strategy equilibrium of the mechanism if for all i and all θ_i

$$u_i(g(s_i^*(\theta_i), s_{-i}), \theta_i) \geq u_i(g(s_i', s_{-i}), \theta_i) \quad \text{for all } s_i' \in S_i \text{ and all } s_{-i} \in S_{-i}.$$

Definition

The mechanism $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ implements the social choice function $f(\cdot)$ in dominant strategies if there exists a dominant strategy equilibrium of Γ, $s^*(\cdot)$ such that

$$g(s^*(\theta)) = f(\theta) \qquad \text{for all } \theta \in \Theta.$$

Direct Mechanisms and Revelation Principle

- As before, we can restrict ourselves (without loss of generality) to direct mechanisms, in which $S_i = \Theta_i$ for all i and $g(\theta) = f(\theta)$ for all θ.

Definition

The social choice function $f(\cdot)$ is truthfully implementable in dominant strategies (or dominant strategy incentive compatible, or strategy-proof) if the direct mechanism $\Gamma = (\Theta_1, \ldots, \Theta_I, f(\cdot))$ has a dominant strategy equilibrium $s^*(\cdot)$ such that $s_i^*(\theta_i) = \theta_i$ for all θ_i, and all i, i.e., for all i and θ_i,

$$u_i(f(\theta_i, \theta_{-i}), \theta_i) \geq u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i) \quad \textit{for all } \hat{\theta}_i, \theta_{-i}.$$

Proposition (The Revelation Principle for Dominant Strategies)

Suppose that there exists a mechanism $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ that implements the social choice function $f(\cdot)$ in dominant strategies. Then f is truthfully implementable in dominant strategies.

Quasilinear Utilities: Groves-Clarke Mechanisms

- We focus on the special class of environments in which agents have quasilinear utilities. In particular, an alternative y is now a vector $y = (x, t)$, where x can be viewed as an allocation decision and t is the vector of payments.
- Agent i's utility function takes the quasilinear form

$$u_i(y, \theta_i) = v_i(x, \theta_i) + t_i \qquad \text{for all } i = 1, \ldots, I.$$

- A social choice function in this quasilinear environment takes the form $f(\cdot) = (x(\cdot), t_1(\cdot), \ldots, t_I(\cdot))$.
- We are interested in (ex-post) efficient social choice functions:
  - Efficient allocations $x^*(\theta)$ must satisfy for all θ

$$\sum_{i=1}^{I} v_i(x^*(\theta), \theta_i) \geq \sum_{i=1}^{I} v_i(x, \theta_i) \qquad \text{for all } x \in X,$$

$x^*(\theta)$ maximizes the social utility.

Truthful Implementation of Efficient Allocations

Proposition (Groves (73))

The social choice function $f(\cdot) = (x^*(\cdot), t_1(\cdot), \ldots, t_I(\cdot))$ *is truthfully implementable in dominant strategies if for all* i,

$$t_i(\theta) = \left[\sum_{j \neq i} v_j(x^*(\theta), \theta_j)\right] + h_i(\theta_{-i}), \tag{1}$$

where $h_i(\cdot)$ *is an arbitrary function of* θ_{-i}.

Proof: Suppose truth telling is not a dominant strategy for some i, i.e., there exists some $\theta_i, \hat{\theta}_i$ and θ_{-i} such that

$$v_i(x^*(\hat{\theta}_i, \theta_{-i}), \theta_i) + t_i(\hat{\theta}_i, \theta_{-i}) > v_i(x^*(\theta_i, \theta_{-i}), \theta_i) + t_i(\theta_i, \theta_{-i}).$$

Substituting from Eq. (1), this yields

$$\sum_{j=1}^{I} v_j(x^*(\hat{\theta}_i, \theta_{-i}), \theta_j) > \sum_{j=1}^{I} v_j(x^*(\theta), \theta_j), \qquad \text{contradiction.}$$

- *Intuition:* The transfer depends on his announced type only through the allocation rule. The change in i's transfer reflects exactly the externality he is imposing on other agents.

Special Case: Clarke Mechanism

- A special case of a Groves mechanism was discovered independently by Clarke and is known as the Clarke mechanism.
- This mechanism corresponds to the case in which

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

where $x^*_{-i}(\theta_{-i})$ satisfies

$$\sum_{j \neq i} v_j(x^*_{-i}(\theta_{-i}), \theta_j) \geq \sum_{j \neq i} v_j(x, \theta_j) \qquad \text{for all } x \in X, \text{and all } \theta_{-i}.$$

- Hence, $x^*_{-i}(\cdot)$ is the efficient allocation with $I - 1$ agents.

Transfer in Clarke Mechanism

- Agent i's transfer in the Clarke mechanism is given by

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

Remarks:

- The transfer satisfies $t_i(\theta) = 0$ if agent i's announcement does not change the allocation decision, and $t_i(\theta) < 0$ if it does; i.e. if he is "pivotal" to the efficient allocation. This implies that agent i needs to pay the amount of resource he uses or the damage he imposes on the system and others.
- Consider allocation of a single indivisible good. In this case *Clarke mechanism* implements the social choice function implemented by second price auction (or the Vickrey auction):
  - $x^*(\theta)$: allocate to the highest valuation buyer.
  - "pivotal" when he has the highest valuation
  - when pivotal, the first term in the transfer expression is equal to 0, and the second term gives the second highest valuation.
- Hence, these mechanisms are commonly referred to as the Vickrey-Clarke-Groves (VCG) mechanisms.

Uniqueness of Groves Payments for Efficiency

- We have seen so far that under Groves payments, we can implement efficient allocations.
- The question then arises: are these the only social choice function that achieves an efficient allocation?
- The answer is yes under some mild conditions, establish by Green and Laffont.

Proposition (Green and Laffont (79))

Let $\mathcal{V}$ denote the set of all functions $v : X \to \mathbb{R}$. Suppose for all i, we have $\{v_i(\cdot, \theta_i) \mid \theta_i \in \Theta_i\} = \mathcal{V}$ (i.e., every possible valuation function arises for some θ_i). Then, a social choice function $f(.) = (X^*(.), t(.))$ with an efficient allocation is truthfully implemented in dominant strategies only if $t_i(\cdot)$ is given by the Groves payments.

Proof

- Note that for all θ, we can write

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

- We will show that if $f(\cdot)$ is strategy proof, then $h_i(\cdot)$ must be independent of θ_i.
- Suppose it is not, i.e., there exists $\theta_i, \hat{\theta}_i, \theta_{-i}$ such that, $h_i(\theta_i, \theta_{-i}) \neq h_i(\hat{\theta}_i, \theta_{-i})$.
- We consider two cases:
- **Case 1:** $x^*(\theta_i, \theta_{-i}) = x^*(\hat{\theta}_i, \theta_{-i})$: By IC in dominant strategies, we have

$$\begin{aligned} v_i(x^*(\theta_i, \theta_{-i}), \theta_i) + t_i(\theta_i, \theta{-}i) &\geq v_i(x^*(\hat{\theta}_i, \theta_{-i}), \theta_i) + t_i(\hat{\theta}_i, \theta{-}i) \\ v_i(x^*(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i) + t_i(\hat{\theta}_i, \theta{-}i) &\geq v_i(x^*(\theta_i, \theta_{-i}), \hat{\theta}_i) + t_i(\theta_i, \theta_{-i}), \end{aligned}$$

implying that $t_i(\theta_i, \theta_{-i}) = t_i(\hat{\theta}_i, \theta_{-i})$, and therefore $h_i(\theta_i, \theta_{-i}) = h_i(\hat{\theta}_i, \theta_{-i})$, contradiction.

Proof

- **Case 2:** $x^*(\theta_i, \theta_{-i}) \neq x^*(\hat{\theta}_i, \theta_{-i})$: Suppose without loss of generality that $h_i(\theta_i, \theta_{-i}) > h_i(\hat{\theta}_i, \theta_{-i})$.
- Define type θ_i^ϵ for some $\epsilon > 0$ as:

$$\begin{aligned} v_i(x, \theta_i^\epsilon) \quad = \quad & -\sum_{j \neq i} v_j(x^*(\theta_i, \theta_{-i}), \theta_j) \text{ if } x = x^*(\theta_i, \theta_{-i}) \\ & -\sum_{j \neq i} v_j(x^*(\hat{\theta}_i, \theta_{-i}), \theta_j) + \epsilon \text{ if } x = x^*(\theta_i, \theta_{-i}) \\ & -\infty \text{ otherwise} \end{aligned}$$

- We show for sufficiently small ϵ, type θ_i^ϵ will report θ_i (when other report θ_{-i}).
- Note that $x^*(\theta_i^\epsilon, \theta_{-i}) = x^*(\hat{\theta}_i, \theta_{-i})$, since $x^*(\hat{\theta}_i, \theta - i)$ maximizes $v_i(x, \theta_i^\epsilon) + \sum_{j \neq i} v_j(x, \theta_j)$.
- Now truth telling being a dominant strategy implies

$$v_i(x^*(\hat{\theta}_i, \theta_{-i}), \theta_i^\epsilon) + t_i(\theta_i^\epsilon, \theta_{-i}) \geq V_i(X^*(\theta_i, \theta_{-i}), \theta_i^\epsilon) + t_i(\theta_i, \theta_{-i}).$$

Proof

- Substituting yields
$$\epsilon + h_i(\theta_i^\epsilon, \theta_{-i}) \geq h_i(\theta_i, \theta_{-i}).$$
- But since, $x_i^*(\theta_i^\epsilon, \theta_{-i}) = x_i^*(\hat{\theta}_i, \theta_{-i})$, we get $h_i(\theta_i^\epsilon, \theta_{-i}) = h_i(\hat{\theta}_i, \theta_{-i})$ (by part (1)).
- This implies $\epsilon + h_i(\hat{\theta}_i, \theta_{-i}) \geq h_i(\theta_i, \theta_{-i})$, which together with $h_i(\theta_i, \theta_{-i}) > h_i(\hat{\theta}_i, \theta_{-i})$ yields a contradiction for sufficiently small ϵ.

Budget-Balancedness

- So far, we have studied whether we can implement in dominant strategies a social choice function that results in an efficient allocation.
- Another property that ex-post efficiency requires is the budget balance condition:

$$\sum_i t_i(\theta) = 0 \qquad \text{for all } \theta,$$

  i.e., there are no net transfers in or out of the system.
- Unfortunately, in many cases (i.e., if the set of possible types for each agent sufficiently rich), it is impossible to truthfully implement efficient allocations and budget balancedness in dominant strategies.

Proposition (Green and Laffont/Hurwicz)

There is no social choice function that is truthfully implementable in dominant strategies and is efficient (allocationwise) and budget-balanced.

- Therefore, as a next step we relax the dominant strategy implementation.

Implementation in Bayesian Nash Equilibrium

Definition

The strategy profile $s^*(\cdot) = (s_1^*(\cdot), \ldots, s_I^*(\cdot))$ is a Bayesian Nash equilibrium of the mechanism $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ if for all i and for all θ_i,

$$E_{\theta_{-i}}[u_i(g(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i)|\theta_i] \geq E_{\theta_{-i}}[u_i(g(\hat{s}_i, s_{-i}^*(\theta_{-i})), \theta_i)|\theta_i],$$

for all $\hat{s}_i \in S_i$.

Definition

The mechanism Γ implements the social choice function $f(\cdot)$ in Bayesian Nash equilibrium if there is a Bayesian Nash equilibrium of Γ, $s^*(\cdot)$, such that $g(s^*(\theta)) = f(\theta)$ for all θ.

Implementation in Bayesian Nash Equilibrium

Definition

The social choice function $f(\cdot)$ *is truthfully implementable in Bayesian Nash equilibrium (or is Bayesian incentive compatible) if* $s_i^*(\theta_i) = \theta_i$ *for all* θ_i *and all* i *is a Bayesian Nash equilibrium of the direct revelation mechanism* $\Gamma = (\Theta_1, \ldots, \Theta_I, f(\cdot))$*, i.e., for all* i *and* θ_i*,*

$$E_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i}), \theta_i)|\theta_i] \geq E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i)|\theta_i],$$

for all $\hat{\theta}_i \in \Theta_i$.

- Revelation principle principle again allows us to restrict attention to truthful equilibria of direct mechanisms.

Expected Externality Mechanism

- We will now show that it is possible to implement efficient (allocation-wise) and budget-balanced social choice functions in Bayesian Nash equilibrium.
- This mechanism is due to d'Aspremont and Gerard-Varet (hence the abbreviation dAGV mechanisms).
- Let $x^*(\cdot)$ be the efficient allocation as defined before and let

$$t_i(\theta) = E_{\tilde{\theta}_{-i}}\Big[\sum_{j \neq i} v_j(x^*(\theta_i, \tilde{\theta}_{-i}), \tilde{\theta}_j)\Big] + h_i(\theta_{-i}).$$

- We can show that these transfers together with efficient allocations satisfy both Bayesian incentive compatibility and budget-balancedness.