INSE6441 Jia Yuan Yu

8: Revelation Principle, VCG Mechanism

Concordia March 10, 2017

1 Mechanism Design

We have decision makers with private types, and hence private utility functions. We want to make a collective decision based on these private utility functions (e.g., resource allocation, matching, etc.). To take the correct decision, we need to find out every utility function. However, individual decision makers may act strategically by misreporting their private utilities in order to improve their payoff. This is the mechanism design problem.

Consider a Bayesian game with players $1, \ldots, I$, strategy spaces $S_1, \ldots, S_I$,type spaces $T_1, \ldots, T_I$, utility functions $u_1, \ldots, u_I$. Let $S = S_1 \times \ldots \times S_I$.

Definition 1.1 (Social choice function)**.** A social choice function $f : T_1 \times \ldots \times T_I \to S$ maps a type profile $(\theta_1, \ldots, \theta_I)$ to an outcome (collective decision) $f(\theta_1, \ldots, \theta_I) \in S$.

In order to determine $f(\theta_1, \ldots, \theta_I)$, we unfortunately need to know $\theta_1, \ldots, \theta_I$. Some players may increase their utility by misreporting their type.

Definition 1.2 (Mechanism)**.** A mechanism is a function $m : T \to S$ mapping the reported type profile to an outcome.

Definition 1.3 (Implementability)**.** A social choice function f is implementable if there exists a Bayesian game with equilibrium strategy profile $(s_1^*, \ldots, s_I^*)$ such that for every type profile $t \in T$, the outcome coincides with the social choice function:

$$(s_1^*(t_1), \ldots, s_I^*(t_I)) = f(t_1, \ldots, t_I).$$

Definition 1.4 (Truthfully implementable, incentive compatible, strategy-proof)**.** A social choice function f is truthfully implementable—or incentive compatible, if there exists a Bayesian game with equilibrium strategy profile $(s_1^*, \ldots, s_I^*)$ such that for every player i and type $t_i \in T_i$:

$$s_i^*(t_i) = t_i.$$

These two concepts are related.

Theorem 1.1 (Revelation Principle)**.** *If a social choice function f is implementable, then f is also truthfully implementable.*

Proof. If f is implementable, then (by the definition of Bayesian equilibrium) for all i and $z \in S_i$:

$$\mathbb{E}u_i((s_i^*(t_i), s_{-i}(\mathbf{t}_{-i})), t_i, \mathbf{t}_{-i}) \geq \mathbb{E}u_i((z, s_{-i}(\mathbf{t}_{-i})), t_i, \mathbf{t}_{-i}).$$

This implies that for all i and $w \in T_i$:

$$\mathbb{E}u_i((s_i^*(t_i), s_{-i}(\mathbf{t}_{-i})), t_i, \mathbf{t}_{-i}) \geq \mathbb{E}u_i((s_i^*(w), s_{-i}(\mathbf{t}_{-i})), t_i, \mathbf{t}_{-i}).$$

In turn, by the definition of implementability, we have for all i and $w \in T_i$:

$$\mathbb{E}u_i(f(t_i, \mathbf{t}_{-i}), t_i, \mathbf{t}_{-i}) \geq \mathbb{E}u_i(f(w, \mathbf{t}_{-i}), t_i, \mathbf{t}_{-i}).$$

Hence, $(t_1, \ldots, t_I)$ is a Bayesian equilibrium of the corresponding game with utility function $w_i(s_i, s_{-i}, t_i, t_{-i}) = u_i(f(s_i, s_{-i}), t_i, t_{-i})$. □

Remark 1. The revelation principle says that if we can define a game that implements a social choice function f, then we can define a new game where each player i has a proxy that plays the equilibrium strategy $s_i(t_i)$ for that player, then telling the truth is an equilibrium in this new game.

Remark 2. The first-price auction is truthfully implementable if every player reports their valuation v_i and a proxy bids according to $\pi(v_i)$.

2 Allocation of goods optimizing social welfare

Consider a resource allocation problem with I players, with types $t_1 \in T_1, \ldots, t_I \in T_I$, and let S denote the set of outcomes (possible allocations of a set of goods). Observe that the sets S and T_i can be huge. The payoff to player i for outcome $s \in S$ has the following quasilinear form:

$$u_i(s, t) = v(s, t_i) - p_i(s),$$

where v denotes a parametrized function (parametrized by the type t_i) describing player i's valuation for an outcome s, and p_i is the payment from player i. Can we implement a (utilitarian in the sense of Jeremy Bentham) social choice function that maximizes $\sum_i v(s, t_i)$? Yes!

By the revelation principle, we can look for a direct revelation game where players report their types. Consider the following mechanism (due to Vickrey, Clarke, Groves). Each player i reports $\hat{t}_i$, a central agent computes

$$s^*(\hat{t}) = \arg\max_{s \in S} \sum_{i=1}^{I} v(s, \hat{t}_i),$$

and each player i pays

$$p_i(\hat{t}) = \underbrace{\left(\max_{s \in S} \sum_{j \neq i} v(s, \hat{t}_j)\right)}_{\text{others' welfare without } i} - \underbrace{\sum_{j \neq i} v(s^*(\hat{t}), \hat{t}_j)}_{\text{others' welfare with } i}.$$

This payment scheme (Clarke's pivot rule) corresponds to player i's effect on the social welfare—or externality. Observe that this payment is nonpositive.

Check that reporting $\hat{t}_i = t_i$ is an equilibrium (by dominance). The strategy profile is $(t_1, \dots, t_I)$, the corresponding outcome is $s^*(t)$, we have

$$\begin{aligned} u_i(s^*(t), t) &= v(s^*(t), t_i) - p_i(s^*(t)) \\ &= \underbrace{v(s^*(t), t_i) + \sum_{j \neq i} v(s^*(t), t_j)}_{\text{maximized already by } s^*(t)} - \left(\max_s \sum_{j \neq i} v(s, t_j) \right) \\ &\geq v(s^*(z, t_{-i}), t_i) + \sum_{j \neq i} v(s^*(z, t_{-i}), t_j) - \left(\max_s \sum_{j \neq i} v(s, t_j) \right), \end{aligned}$$

for all $z \in T_i$.

Example 2.1 (Vickrey auction)**.** The Vickrey or second-price auction is an instance of the VCG mechanism above. There are I outcomes corresponding to the possible winners. The value of player i for each outcome is 0 or v_i. Finding $s^*(\hat{t})$ corresponds to finding the highest bidder. The payment is the second highest bid.

Example 2.2 (VCG auction)**.** Consider M items (or online ad slots) and I buyers. The set S is the set of all possible allocations of these m items among the buyers (there are I^M such allocations!). The type t_i of player i describes the value $v(x, t_i)$ derived from every subset $x \subseteq \{1, \dots, M\}$. This problem is hard to solve.

For a simpler version of this problem, suppose that each player does not get more payoff from obtaining more than one item. In this case, computing $s^*(t)$ consists of solving a weighted bipartite matching problem (the weights correspond to the values of item j to buyer i). It can be solve by the Hungarian algorithm. If the weights are all the same, then it can be solved by the Ford-Fulkerson algorithm (cf. max-flow min-cut problem, supply chain course).

Example 2.3 (Funding public project)**.** Installing internet costs 100. Two players, valuations v_1 and v_2, two outcomes: install or not. Each player first contributes 50 into a pot. The players submit $\hat{v}_1, \hat{v}_2$ to the VCG mechanism, which computes for player 1 the net utilities (after equal contribution of 50):

$$\begin{aligned} v(\text{install}, \hat{v}_1) &= \hat{v}_1 - 50, \\ v(\text{not install}, \hat{v}_1) &= 50 - 50 = 0, \end{aligned}$$

for player 2:

$$\begin{aligned} v(\text{install}, \hat{v}_2) &= \hat{v}_2 - 50, \\ v(\text{not install}, \hat{v}_2) &= 0, \end{aligned}$$

the outcome:

$$\arg\max_{z=\text{install, not install}} v(z,\hat{v}_1)+v(z,\hat{v}_2)=\begin{cases}\text{install} & \text{if } \hat{v}_1+\hat{v}_2-100\geq 0,\\ \text{not install} & \text{otherwise,}\end{cases},$$

payments (in addition to 50):

$$p_1(\hat{v})=\begin{cases}\hat{v}_2-50-(\hat{v}_2-50) & \text{if } \hat{v}_2\geq 100,\\ 0-(\hat{v}_2-50) & \text{if } \hat{v}_2<100, \hat{v}_1+\hat{v}_2\geq 100.\\ 0-0 & \text{if } \hat{v}_1+\hat{v}_2<100.\end{cases}$$

$$p_2(\hat{v})=\begin{cases}0 & \text{if } \hat{v}_1\geq 100,\\ 50-\hat{v}_1 & \text{if } \hat{v}_1<100, \hat{v}_1+\hat{v}_2\geq 100.\\ 0 & \text{if } \hat{v}_1+\hat{v}_2<100.\end{cases}$$

However, observe that the total payment is $50+50+p_1(\hat{v})+p_2(\hat{v})$, which may exceed the cost 100 of the public good.

Example 2.4 (Bilateral trade)**.** The seller has an item. Buyer valuation of the item is $v_b\in[0,1]$, seller's valuation is $v_s\in[0,1]$. Outcomes: trade, no trade. The players submit $\hat{v}_b,\hat{v}_s$ to the VCG mechanism, which computes for the seller:

$$v(\text{trade},\hat{v}_s)=-\hat{v}_s,$$
$$v(\text{no trade},\hat{v}_s)=0,$$

for the buyer:

$$v(\text{trade},\hat{v}_b)=\hat{v}_b,$$
$$v(\text{no trade},\hat{v}_b)=0,$$

the outcome:

$$\arg\max_{z=\text{trade, no trade}} v(z,\hat{v}_s)+v(z,\hat{v}_b)=\begin{cases}\text{trade} & \text{if } \hat{v}_b-\hat{v}_s\geq 0,\\ \text{no trade} & \text{otherwise,}\end{cases},$$

and the payments:

$$p_b(\hat{v})=\begin{cases}0 & \text{if } \hat{v}_b\leq\hat{v}_s,\\ v_s & \text{if } \hat{v}_b>\hat{v}_s\end{cases}$$

$$p_s(\hat{v})=\begin{cases}0 & \text{if } \hat{v}_b\leq\hat{v}_s,\\ -v_b & \text{if } \hat{v}_b>\hat{v}_s\end{cases}$$

Finally, observe that the VCG mechanism has a nice property: if $v(s, t_i) \geq 0$ for all s, then

$$
\begin{aligned}
u_i(s^*(t), t) &= v(s^*(t), t_i) + \sum_{j \neq i} v(s^*(t), t_j) - \left(\max_s \sum_{j \neq i} v(s, t_j) \right) \\
&= \left(\max_{s' \in S} \sum_{i=1}^{I} v(s', t_i) \right) - \underbrace{\left(\max_s \sum_{j \neq i} v(s, t_j) \right)}_{\text{Let } s^+ \text{ denote the arg max}} \\
&\geq \left(\max_{s' \in S} \sum_{i=1}^{I} v(s', t_i) \right) - \left(\sum_{j \neq i} v(s^+, t_j) + v(s^+, t_i) \right) \\
&= \left(\max_{s' \in S} \sum_{i=1}^{I} v(s', t_i) \right) - \left(\sum_{i=1}^{I} v(s^+, t_i) \right) \geq 0.
\end{aligned}
$$

This means that a rational player would not mind participating in the game.

3 Reading material

- Chapter 7 of Fudenberg and Tirole.
- Chapter 6 of Myerson.
- Chapter 23 of Mas-Colell, Whinston, Green.