INSE6441 Jia Yuan Yu

7: Mechanism Design

Concordia March 7, 2017

So far, we have considered the setting where someone else fixes the game to play, and we have to predict how participants will interact. In this and the following lectures, we consider the converse problem: given a desired outcome, how should we design the rules of interaction such that the desired outcome arises naturally from the interactions of rational decision makers.

For example, given that we want to give an item to the person who most values it, how should we proceed? Auction. Given that King Solomon wants to give a baby to its true mother, how should he proceed? Given that the government wants to construct a public good that benefits everyone, how should it levy the fees?

Example 0.1 (First-price auction)**.** Each student generates a random number uniformly on $[0, 10]$ for their valuation, and bids according to their valuation. Winner pays bid and gets their valuation.

1 Another example of Bayesian games

Example 1.1 (First-price auction as Bayesian game)**.** Real-life auctions are also better modeled as Bayesian games since opponents' valuations are known only up to a probability distribution. Consider a first-price auction with I bidders (Section 3.11 of Myerson). The valuations $V_1, \ldots, V_I$ of the bidders are their types and modeled by i.i.d. random variables with probability distribution F taking values over the support $[0, M]$.

For fixed bid profile b and type profile v, the payoff for bidder i is

$$u_i(b, v) = \begin{cases} v_i - b_i & \text{if } b_i > \max_{j \neq i} b_j, \\ 0 & \text{otherwise.} \end{cases}$$

Let's look for a symmetric Bayesian equilibrium $(\pi_1, \ldots, \pi_I)$, where the strategies satisfy $\pi_1 = \ldots = \pi_I = \pi$, where $\pi : [0, M] \to [0, M]$. In this setting, the valuations are independent, hence $\mathbb{P}(V_j < z \mid V_i = v_i) = F(z)$. Moreover, the types have a continuous distribution. Hence, the objective function of i in the Bayesian game becomes

$$\begin{aligned} &\int_{v_1 \in [0,\infty)} \ldots \int_{v_{i-1} \in [0,\infty)} \int_{v_{i+1} \in [0,\infty)} \ldots \int_{v_I \in [0,\infty)} u_i(\pi_1(v_1), \ldots, \pi_{i-1}(v_{i-1}), z, \pi_{i-1}(v_{i-1}), \pi_I(v_I), v) \\ &\qquad \mathrm{d}F(v_1) \ldots \mathrm{d}F(v_{i-1}) \mathrm{d}F(v_{i+1}) \ldots \mathrm{d}F(v_I) \\ &= \int_{v_1 : \pi(v_1) < z} \ldots \int_{v_{i-1} :: \pi(v_{i-1}) < z} \int_{v_{i+1} : \pi(v_{i+1}) < z} \ldots \int_{v_I : \pi(v_I) < z} (v_i - z) \\ &\qquad \mathrm{d}F(v_1) \ldots \mathrm{d}F(v_{i-1}) \mathrm{d}F(v_{i+1}) \ldots \mathrm{d}F(v_I) \\ &= (v_i - z) F^{I-1}(\pi^{-1}(z)), \end{aligned}$$

where we observed that the integral is nonzero if $\pi(v_j) < z$ or $v_j < \pi^{-1}(z)$ for all $j \neq i$.

The Bayesian equilibrium strategy π must therefore satisfy[1]

$$\pi(v_i) \in \arg\max_z \quad (v_i - z)F^{I-1}(\pi^{-1}(z)), \quad \text{for all } v_i \in [0, M].$$

Fix an arbitrary v_i. By setting the derivative[2] of the objective to zero, each player's best response z^* must satisfy:

$$(v_i - z^*)(I-1)F^{I-2}(\pi^{-1}(z^*))F'(\pi^{-1}(z^*))\frac{1}{\pi'(\pi^{-1}(z^*))} - F^{I-1}(\pi^{-1}(z^*)) = 0.$$

Since $z^* = \pi(v_i)$, the above equation requires that

$$\begin{aligned}(I-1)(v_i - \pi(v_i))F^{I-2}(v_i)F'(v_i)\frac{1}{\pi'(v_i)} &= F^{I-1}(v_i)\\ (I-1)(v_i - \pi(v_i))F'(v_i) &= \pi'(v_i)F(v_i)\\ \pi'(v_i) &= (I-1)(v_i - \pi(v_i))\frac{F'(v_i)}{F(v_i)}.\end{aligned}$$

To solve this differential equation for π, try:

$$\begin{aligned}\pi(x) &= x - \frac{1}{F^{I-1}(x)}\int_0^x F^{I-1}(z)\mathrm{d}z,\\ \pi'(x) &= 1 - \left(-(I-1)F^{-I}(x)F'(x)\int_0^x F^{I-1}(z)\mathrm{d}z + 1\right).\end{aligned}$$

Example 1.2 (First-price auction with uniformly distributed valuations)**.** Suppose that $v_1, \ldots, v_I$ are i.i.d. with uniform distribution on $[0, 1]$. The equilibrium strategy is

$$\pi(v) = v - \frac{1}{v^{I-1}}\int_0^v z^{I-1}\mathrm{d}z = v - \frac{1}{v^{I-1}}\frac{v^I}{I} = \frac{I-1}{I}v.$$

Remark 1. Tenders for government contracts are first-price auctions.

2 Social Choice

By desired outcome, we mean making a decision in a group (division of labor in a project, allocation of goods)—social or collective decision making, or aggregation of individual preferences into a social choice. In this setting, we don't consider strategic behaviour in the individual decision makers.

Example 2.1. Voting on two possible midterm dates. Majority outcome is unique.

[1] We have to assume that π is invertible.

[2] Use the chain rule, and the fact that $\frac{\mathrm{d}}{\mathrm{d}z}\pi^{-1}(z) = \frac{1}{\pi'(\pi^{-1}(z))}$. We also have to assume that π is differentiable.

Example 2.2 (Condorcet paradox)**.** Voting on three midterm dates is tricky! Pairwise majority preferences lead to a cycle.

$$\begin{aligned} &\text{Voter 1} \quad A \succ B \succ C \\ &\text{Voter 2} \quad B \succ C \succ A \\ &\text{Voter 3} \quad C \succ A \succ B \end{aligned}$$

3 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 3.1 (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.

Example 3.1 (Misrepresent type, Example 23.B.1 of MWG)**.** Two players. Player 1 has only one type $T_1 = \{\theta_1'\}$. Player 2 has two possible types $T_2 = \{\theta_2', \theta_2'', \theta_2'''\}$. The players are asked to report their individual types, the action spaces are $S_1 = \{\theta_1'\}$ and $S_2 = \{\theta_2', \theta_2''\}$. For simplicity, let's relabel the set of outcomes: $S = \{x, y, z\}$. The utility functions are:

$$\text{type } \theta_1' : \quad u_1(x, \theta_1') > u_1(y, \theta_1') > u_1(z, \theta_1').$$

$$\text{type } \theta_2' : \quad u_2(z, \theta_2') > u_2(y, \theta_2') > u_2(x, \theta_2'),$$

$$\text{type } \theta_2'' : \quad u_2(y, \theta_2'') > u_2(x, \theta_2'') > u_2(z, \theta_2''), \text{type } \theta_2''' : \quad u_2(y, \theta_2'') > u_2(x, \theta_2'') > u_2(z, \theta_2''),$$

Suppose that we want to implement a social choice function taking the following values:

$$\begin{aligned} f(\theta_1', \theta_2') &= y, \\ f(\theta_1', \theta_2'') &= x. \end{aligned}$$

Unfortunately, since we are in a game of incomplete information, we can only evaluate $f(s_1(t_1), s_2(t_2))$. Even if $t_2 = \theta_2''$, player 2 can improve its utility by reporting $s_2(\theta_2'') = \theta_2'$ instead of $s_2(\theta_2'') = \theta_2''$.

Example 3.2 (Auctions, Example 23.B.4 of MWG)**.** The players are $0, 1, \ldots, I$, whose types are their valuations $t = v = (v_0, v_1, \ldots, v_I)$. Player 0 is the seller, assume that $v_0 = 0$. For simplicity, we relabel the outcomes, letting $y_i = 1$ only if player i receives the good (0 otherwise), τ_i denote the money received by player i (negative if the player pays a positive amount):

$$S = \left\{ (y_0, \ldots, y_I, \tau_0, \ldots, \tau_I) : y_i \in \{0, 1\}, \tau_i \in \mathbb{R}, \text{ for } i = 0, \ldots, I, \sum_i y_i = 1, \sum_i \tau_i = 0 \right\}.$$

Player i's utility function is

$$u_i(y, \tau, v) = v_i y_i + \tau_i.$$

A social choice function f (a mapping from valuation profile to outcome) for this problem is efficient if the outcome $f(v) = (y_1(v), \ldots, y_I(v), \tau_1(v), \ldots, \tau_I(v))$ allocates the good to a player with the highest valuation:

$$y_i(v)(v_i - \max\{v_1, \ldots, v_I\}) = 0, \quad \text{for all } i$$

and money is conserved (transferred with no waste)

$$\sum_i \tau_i(v) = 0.$$

Suppose that there are two buyers ($I = 2$), that the valuations V_1, V_2 are independent random variables with the uniform distribution p on $[0, 1]$, and that the buyers have the same belief $p_1(V_2 \leq z) = p_2(V_1 \leq z)$ for all z, and this belief is consistent with the actual distribution p. The outcome $\hat{f}(v) = (\hat{y}(v), \hat{\tau}(v))$ corresponding to first-price auction is

$$\begin{aligned}
\hat{y}_0(v) &= 0, \quad \text{for all } v, \\
\hat{y}_1(v) &= \begin{cases} 1 & \text{if } v_1 \geq v_2, \\ 0 & \text{otherwise.} \end{cases} \\
\hat{y}_2(v) &= \begin{cases} 1 & \text{if } v_2 > v_1, \\ 0 & \text{otherwise.} \end{cases} \\
\hat{\tau}_1(v) &= -v_1 y_1(v), \\
\hat{\tau}_2(v) &= -v_2 y_2(v), \\
\hat{\tau}_0(v) &= -\hat{\tau}_1(v) - \hat{\tau}_2(v) = v_1 y_1(v) + v_2 y_2(v).
\end{aligned}$$

The buyer with the highest valuation receives the good. There is no money wasted in transfer. Hence this outcome is efficient. Suppose that instead of asking for bids from the buyers, we ask them to reveal their valuations (types) v_1, v_2 in order to implement the outcome $(\hat{y}(v), \hat{\tau}(v))$. However, we can only implement the outcome $(\hat{y}(s), \hat{\tau}(s))$,

based on the revealed valuations $s = (s_1, s_2)$. Is the strategy profile $(s_1, s_2) = (v_1, v_2)$ an equilibrium (cf. First-price auction example)? No, because:

$$\begin{aligned} s_1^* &\in \arg\max_z \quad (v_1 - z)\mathbb{P}(V_2 \leq z) \\ &\in \arg\max_z \quad (v_1 - z)z \\ s_1^* &= v_1/2. \\ s_2^* &= v_2/2. \end{aligned}$$

Consider the outcome $\tilde{f}(v) = (\tilde{y}(v), \tilde{\tau}(v))$ corresponding to second-price auction:

$$\begin{aligned} \tilde{y}_0(v) &= 0, \quad \text{for all } v, \\ \tilde{y}_1(v) &= \begin{cases} 1 & \text{if } v_1 \geq v_2, \\ 0 & \text{otherwise.} \end{cases} \\ \tilde{y}_2(v) &= \begin{cases} 1 & \text{if } v_2 > v_1, \\ 0 & \text{otherwise.} \end{cases} \\ \tilde{\tau}_1(v) &= -v_2 y_1(v), \\ \tilde{\tau}_2(v) &= -v_1 y_2(v), \\ \tilde{\tau}_0(v) &= -\tilde{\tau}_1(v) - \tilde{\tau}_2(v) = v_2 y_1(v) + v_1 y_2(v). \end{aligned}$$

This outcome is again efficient. Can we implement the outcome $(\tilde{y}(v), \tilde{\tau}(v))$ by asking the buyers to reveal their types (cf. Second-price auction example, weak dominance)? Yes!

This begs the question: in general, what social choice functions $f : T \to S$ can be implemented in the Bayesian game setting?

3.1 Implementation

Definition 3.2 (Mechanism)**.** A mechanism is a function $m : T \to S$ mapping the reported type profile to an outcome.

Definition 3.3 (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 3.4 (Truthfully implementable, incentive compatible)**.** 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 3.1 (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 $v_i(s_i, s_{-i}, t_i, t_{-i}) = u_i(f(s_i, s_{-i}), t_i, t_{-i})$. □

4 Reading material

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