CS698A: Selected Areas of Mechanism Design **Jan-Apr 2018**

Lecture 1: January 5, 2018

Lecturer: Swaprava Nath *Scribe(s): Harshit Bisht*

This lecture will state definitions and set the ground for the rest of the course.

1.1 Game Theory vs Mechanism Design

Games are strategic interactions among decision making agents that are both **rational** and **intelligent**. When analysing such interactions, game theory takes an agent-based approach, attempting to predict the rational outcome of the game. Mechanism design problems, however, put one in the shoes of the designer attempting to achieve some desired outcome. In this manner, mechanism design is **prescriptive** as opposed to **predictive**.

1.2 An Illustrative Example

Consider the familiar story of when two women approach a king with a baby, each claiming to be the true mother. As the king, setting up an interaction to assign the baby to the true mother becomes a mechanism design problem. The story gives a solution in the form of the king ordering the baby to be cut in half, leading to the true mother giving up her claim for the baby in interests of its safety. However, this solution clearly has no reproducibility, and if she was aware of it, the other woman could have easily reacted the same.

1.3 Definitions and Notation

1.3.1 The Formal Model

All problems considered will be expressed in the following template:

- $N = \{1, ..., n\}$ A finite set of **agents/players** who participate in the mechanism
- X The possible set of **outcomes** that can be realized by the mechanism
- Θ_i Sets of **"types" that the agents can take**, indexed by the list of agents. The type expressed by the agent (not known *a priori*) is denoted by $\theta_i \in \Theta_i$. $\Theta_1 \times ... \times \Theta_n$ is often denoted as simply Θ
- $u_i : X \times \Theta_i \to \mathbb{R}$ **Utility functions** for each agent indexed by the set of agents

1.3.2 Mechanisms

Mechanism design problems always have some goals/objectives related to social welfare in mind while contemplating game design. This will be encapsulated in the **Social Choice Function (SCF)**.

Definition 1.1 *The SCF* $f : \Theta_1 \times ... \times \Theta_n \to X$ *maps the true types of agents to the desirable outcome aligned with some notion of social welfare.*

Since the true types of agents are not revealed to the designer, the goal is to design a decision rule which incentivizes them to reveal their true type in the form of messages.

Definition 1.2 *An (indirect)* ***mechanism*** *is a collection of* ***message spaces*** *indexed by the agents* $M_1, ..., M_n$ *and a* ***decision rule*** $g : M_1 \times ... \times M_n \to X$*.* $M_1 \times ... \times M_n$ *is often denoted as simply* M

Definition 1.3 *A* ***direct mechanism*** *has message spaces as the type spaces* ($M_i = \Theta_i$ *for each agent) and the decision rule coinciding with the social choice function* ($g = f$)

Now that we have defined what mechanisms are, we need to formalize the notion of a "good" mechanism that successfully achieves the desired goal. We define this for direct mechanisms to begin with:

Definition 1.4 *A direct mechanism is said to be* ***truthful/strategy proof/dominant strategy incentive compatible (DSIC)*** *if*

$$u_i(f(\theta_i, \widetilde{\theta_{-i}}), \theta_i) \geq u_i(f(\theta_i', \widetilde{\theta_{-i}}), \theta_i)(\forall \theta_i, \theta_i' \in \Theta_i)(\forall \widetilde{\theta_{-i}} \in \Theta_{-i})\forall i \in N$$

For extending this to indirect mechanisms, we will define what makes a specific message desirable for choice and the conditions needed for truthful messages to be desirable:

Definition 1.5 *In a mechanism* $< M, g >$*, a message* $m_i \in M_i$ *is* ***weakly dominant at*** θ_i *if:*

$$u_i(g(m_i, \widetilde{m_{-i}}), \theta_i) \geq u_i(g(m_i', \widetilde{m_{-i}}), \theta_i)(\forall m_i' \in M_i)\forall \widetilde{m_{-i}} \in M_{-i}$$

Definition 1.6 *A SCF* $f : \Theta \to X$ *is* ***implemented in dominant strategies*** *by* $< M, g >$ *if:*

1. *There exist message mappings* $m_i : \Theta_i \to M_i$ *such that* $m_i(\theta_i)$ *is weakly dominant for agent* i*.*
2. $g(m_1(\theta_1), ..., m_n(\theta_n)) = f(\theta_1, ..., \theta_n)$

Such an SCF is then called ***dominant strategy implementable (DSI)***

It is not hard to see that it is much easier to work with direct mechanisms as opposed to indirect. Fortunately, we can do this without loss of generality because of the following result:

Theorem 1.7 *(Without Proof)* ***Revelation Principle****: If a Social Choice Function is dominant strategy implementable (DSI), then it is also dominant strategy incentive compatible (DSIC)*

1.4 An Illustrative Example: Continued

Let us now model the example discussed in this new language. The two women belong in the set $\mathbf{2} = \{1, 2\}$, with both sharing the type space $\Theta_1 = \Theta_2 = \{"mother", "notmother"\}$. Since the baby is an indivisible good that must go in one piece to one of the agents, this problem is similar to the **Single Object Allocation** problem. Assuming that the true mother values the baby highly, while the other claimant has only a small value for the baby, the king can choose to resolve this via a **Vickrey auction**. The set of outcomes then becomes a collection of (a, p)'s, with a denoting the woman receiving the baby and p the price she pays to the royal court. The SCF in this problem can be modeled as $f("mother", "notmother") = (1, 0)$ and $f("notmother", "mother") = (2, 0)$, with the other f values being inconsequential since we assume atleast one of them is the true mother. The auction would have both message spaces as possible bids $M_1 = M_2 = \mathbb{R}$ and the decision rule as

$$g(p_1, p_2) = \begin{cases} (1, p_2) & \text{if } p_1 > p_2; \\ (2, p_1) & \text{otherwise.} \end{cases}$$

Assuming one of p_1, p_2 is very high while the other is very low, the given decision rule closely approximates the desired SCF.