Classic Mechanism Design

David C. Parkes

Division of Engineering and Applied Science,

Harvard University

CS 286r–Spring 2002

Mechanism Design

Central question: what social choice functions can be implemented in distributed systems with private information and rational agents?

– impose **incentive** based constraints

Note: Mechanism design assumes unlimited computation and communication. Key concept is that of a “rational” agent.

Central Ideas

- **Design criteria**
  - participation constraints
  - incentive-compatibility constraints
  - budget-balance constraints
- **Revelation principle**
- **Impossibility and possibility** results
  - Groves mechanisms
  - median choice mechanisms
  - group-strategy proof cost sharing methods

Basic Definitions

- Set of possible outcomes $\mathcal{O}$
- Agents $i \in \mathcal{I}$, with *preference types* $\theta_i \in \Theta_i$, and $|\mathcal{I}| = N$.
- Utility, $u_i(o, \theta_i)$, over outcome $o \in \mathcal{O}$
- Mechanism $\mathcal{M} = (S, g)$ defines:
  – a *strategy space* $S = S_1 \times \ldots \times S_N$, s.t. agent i chooses a strategy $s_i(\theta_i) \in S_i$, with $s_i : \Theta_i \rightarrow S_i$.
  – an *outcome function*, $g : S^N \rightarrow \mathcal{O}$, s.t. outcome $g(s_1(\theta_1), \ldots, s_N(\theta_N))$ is implemented given strategy profile $s = (s_1(), \ldots, s_N())$.
- *Game:* Utility to agent i from strategy profile s, is $u_i(g(s(\theta)), \theta_i)$, which we write as shorthand $u_i(s, \theta_i)$.

Classic Implementation Concept

Mechanism $\mathcal{M} = (S, g(\cdot))$ **implements** SCF $f(\theta)$ if:

$$g(s_1^*(\theta_1), \ldots, s_N^*(\theta_N)) = f(\theta), \quad \forall \theta \in \Theta^N$$

for an equil. strategy $(s_1^*, \ldots, s_I^*)$; given constraints on the equilibrium concept, transfers, and participation conditions.

Note: the equilibrium concept is left defined, but may be Nash, Bayesian-Nash, or dominant-strategy.

Mechanism Desiderata

- **Efficiency**

  – select the outcome that maximizes total utility

- **Fairness**

  – select the outcome that minimizes the variance in utility

- **Revenue maximization**

  – select the outcome that maximizes revenue to a seller (or more generally, utility to one of the agents)

- **Budget-balanced**

  – implement outcomes that have balanced transfers across agents

- **Pareto Optimal**

  – only implement outcomes o^*, for which for all $o' \neq o^*$, either $u_i(o', \theta_i) = u_i(o^*, \theta_i)$ for all i, *or* $\exists i \in \mathcal{I}$ with $u_i(o', \theta_i) < u_i(o^*, \theta_i)$.

Solution Constraints

- Type of equilibrium (Nash, Bayes-Nash, Dominant, etc.)
- Voluntary participation (“individual rationality”)
- Transfers/no transfers
- Budget-balance

$\Rightarrow$ what can be implemented?

I: Equilibrium Concepts

- **Nash implementation.** Mechanism $\mathcal{M} = (S, g)$ implements $f(\theta)$ in Nash eq. if, for all $\theta \in \Theta$, $g(s^*(\theta)) = f(\theta)$, where $s^*(\theta)$ is a Nash eq., *i.e.*

$$u_i(s_i^*(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i) \geq u_i(s_i'(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i),$$
$$\forall i, \forall \theta, \forall s_i' \neq s_i^*$$

- **Bayes-Nash implementation.** Common prior $F(\theta)$. Mechanism $\mathcal{M} = (S, g)$ implements $f(\theta)$ in Bayes-Nash eq. if, for all $\theta \in \Theta$, $g(s^*(\theta)) = f(\theta)$, where $s^*(\theta)$ is a Bayes-Nash eq., *i.e.*

$$\mathrm{E}_{\theta_{-i}}[u_i(s_i^*(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i)] \geq$$
$$\mathrm{E}_{\theta_{-i}}[u_i(s_i'(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i)],$$
$$\forall i, \forall \theta_i, \forall s_i' \neq s_i^*$$

- **Dominant implementation.** Mechanism $\mathcal{M} = (S, g)$ implements $f(\theta)$ in a dominant strategy eq. if, for all $\theta \in \Theta$, $g(s^*(\theta)) = f(\theta)$, where $s^*(\theta)$ is a dominant strategy eq., *i.e.*

$$u_i(s_i^*(\theta_i), s_{-i}(\hat{\theta}_{-i}), \theta_i) \geq u_i(s_i'(\theta_i), s_{-i}(\hat{\theta}_{-i}), \theta_i),$$
$$\forall i, \theta_i, \hat{\theta}_{-i}, s_i' \neq s_i^*, s_{-i}$$

$$\text{Nash} \supset \text{Bayes-Nash} \supset \text{Dominant}$$

$$\overrightarrow{\text{harder}}$$

II: Participation

Let $\overline{u}_i(\theta_i)$ denote the (expected) utility to agent i with type θ_i of its outside option, and recall that $u_i(f(\theta), \theta_i)$ is the equilibrium utility of agent i from the mechanism.

- *ex ante* individual-rationality; agents choose to participate before they know their own types:

$$E_{\theta \in \Theta}\left[u_i(f(\theta), \theta_i)\right] \geq E_{\theta_i \in \Theta_i} \overline{u}_i(\theta_i)$$

- *interim* individual-rationality
  – agents can withdraw once they know their own type;

$$E_{\theta_{-i} \in \Theta_{-i}}\left[u_i(f(\theta_i, \theta_{-i}), \theta_i)\right] \geq \overline{u}_i(\theta_i)$$

- *ex post* individual-rationality
  – agents can withdraw from the mechanism at the end; $u_i(f(\theta), \theta_i) \geq \overline{u}_i(\theta_i)$.

ex ante $\supset$ interim $\supset$ ex post

$$\overrightarrow{\text{harder}}$$