Bayesian Incentive Compatibility

Types generated by a common prior and are revealed only to the respective agents.

- Recall: Bayesian games

$$\langle N, (M_i)_{i \in N}, (\Theta_i)_{i \in N}, P, (\Gamma_\theta)_{\theta \in \Theta} \rangle$$

$m_i : \Theta_i \to M_i$ message mapping

Definition:

A mechanism $\langle M, g \rangle$ implements SCF f in Bayesian equilibrium if

1. $\exists\, m_i$ s.t. $m_i(\theta_i)$ maximizes the ex-interim expected utility of agent i, ~~$\forall i \in N$~~, $\forall \theta_i \in \Theta_i$, $\forall i \in N$.

i.e.

$$\mathbb{E}_{\theta_{-i}|\theta_i}\left[u_i\left(g(m_i(\theta_i), m_{-i}(\theta_{-i})), \theta_i\right)\right]$$
$$\geq \mathbb{E}_{\theta_{-i}|\theta_i}\left[u_i\left(g(m_i', m_{-i}(\theta_{-i})), \theta_i\right)\right] \quad \forall m_i' \in M_i,\ \forall \theta_i \in \Theta_i,\ \forall i \in N.$$

2. $g(m_i(\theta_i), m_{-i}(\theta_{-i})) = f(\theta_i, \theta_{-i}) \quad \forall \theta \in \Theta.$

Lemma: If an SCF f is implementable in dominant strategies, it is implementable in Bayesian equilibrium.

A direct mechanism $\langle \Theta, f \rangle$ is Bayesian Incentive Compatible (BIC) if (θ_i, θ_{-i}) is a Bayesian equilibrium wrt prior P, i.e.

$$E_{\theta_{-i}|\theta_i}\left[u_i(f(\theta_i, \theta_{-i}), \theta_i)\right] \geq E_{\theta_{-i}|\theta_i}\left[u_i(f(\theta_i', \theta_{-i}), \theta_i)\right]$$

Expectation taken wrt the Bayesian posterior $P(\theta_{-i}|\theta_i)$.

Theorem (Revelation principle for BI SCFs)

If an SCF f is implementable in Bayesian equilibrium, then f is BIC.

Proof: homework.

WLOG always consider IC mechanisms.

A Step Back: Is aggregating opinions easy?

Individual Preferences to Social Preferences

Individual Social Welfare Function Setup:

Arrow's Social Welfare Function Setup:

Finite set of alternatives, $A = \{a_1, \dots, a_m\}$

Finite set of agents, $N = \{1, \dots, n\}$

Each agent has preference ordering over A, denoted by R_i (simpler notation for $\succsim_i$)

$a R_i b$: a is at least as good as b.

Properties of R_i (ordering)

① Completeness: for every pair of alternatives ~~exactly one of the following is true~~: $a, b \in A$, either $a R_i b$ or $b R_i a$

② Reflexivity: $\forall a \in A$, $a R_i a$

③ Transitivity: If $a R_i b$ and $b R_i c \Rightarrow a R_i c$ $\forall a, b, c \in A$, $\forall i \in N$.

Set of all orderings : $\mathcal{R}$

Defn: An ordering R_i is *linear* if for every $a, b \in A$ s.t. $a R_i b$ and $b R_i a \Rightarrow a = b$.

i.e. indifferences are not allowed.

Set of linear orderings $\mathcal{P}$.

R_i → asymmetric part P_i
R_i → symmetric part I_i

Arrovian Social Welfare Function

$F : \mathcal{R}^n \to \mathcal{R}$ domain and range both are rankings

motivation: the collective ordering of the society if one alternative is not feasible, the society can move to the next alternative.

$F(R)$ is an ordering

$\hat{F}(R) \to$ asymmetric part of $F(R)$

$\bar{F}(R) \to$ symmetric part of $F(R)$

Defn. Weak Pareto

An ASWF F satisfies weak pareto if $\forall a, b \in A$ $\forall R \in \mathcal{R}^n$

$$[a P_i b, \forall i \in N] \Rightarrow [a \hat{F}(R) b]$$

Defn. Strong Pareto

An ASWF F satisfies strong Pareto if $\forall a, b \in A, \forall R \in \mathcal{R}^n$

$$[a R_i b, \forall i \in N \text{ and } a P_j b \ \exists j] \Rightarrow [a \hat{F}(R) b]$$

$SP \Rightarrow WP$

We will work with weaker condition WP.