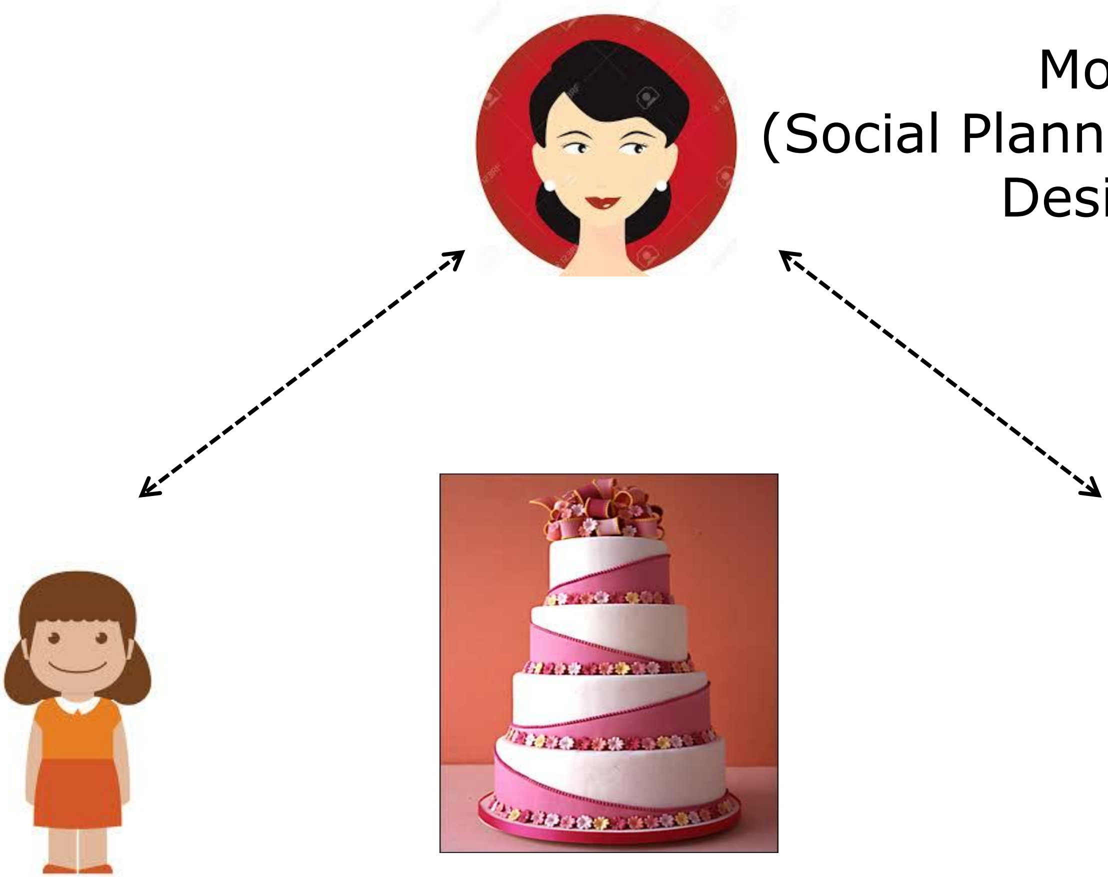
An Introduction to Mechanism Design

Dinesh Garg
IBM India Research Lab, Bangalore
garg.dinesh@in.ibm.com

January 11, 2016

Example 1: One Mother Two Child (Cake Cutting Problem)

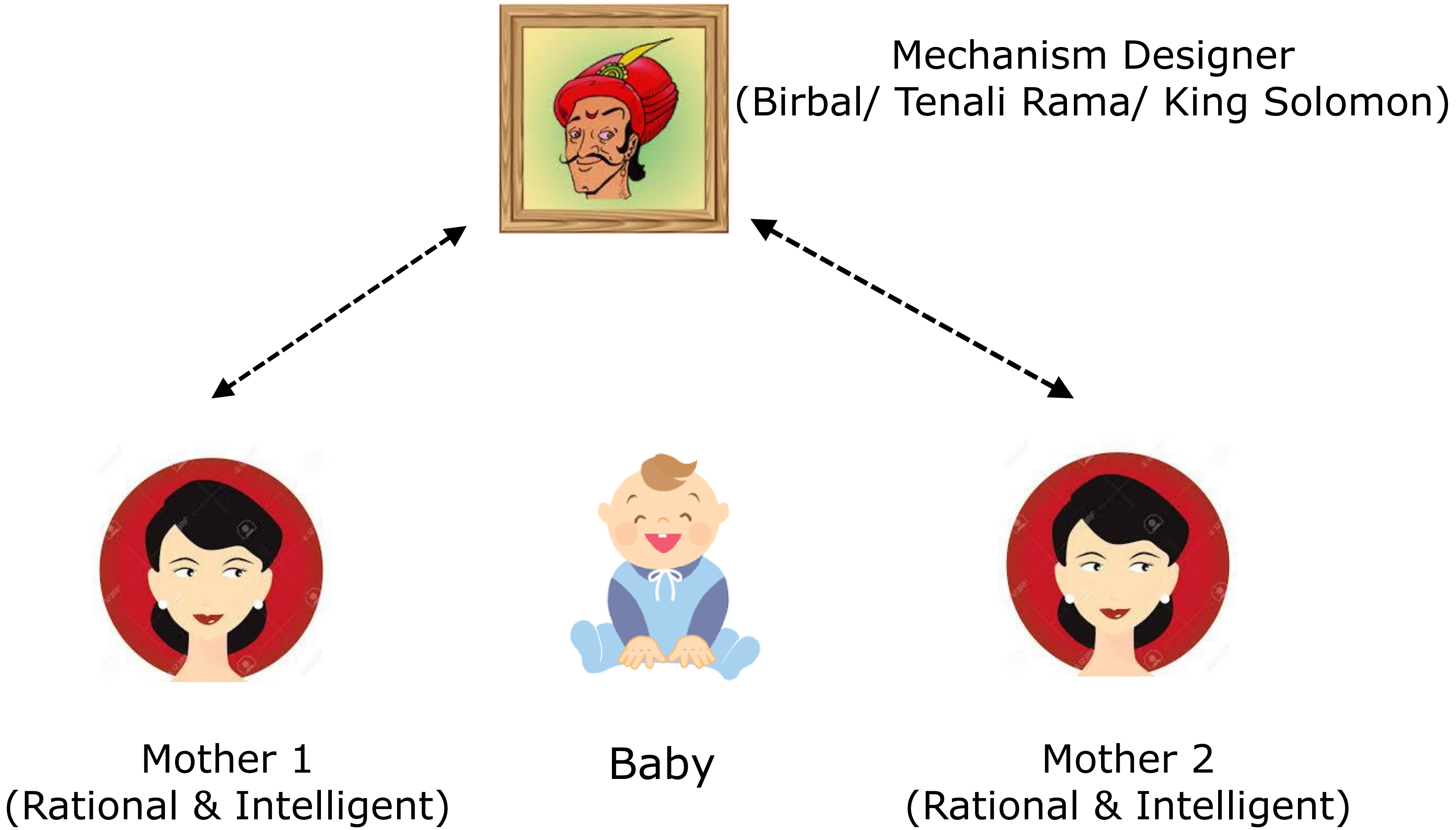


Mother
(Social Planner/ Mechanism Designer)

Children 1
(Rational & Intelligent)

Children 2
(Rational & Intelligent)

Example 2: Two Mothers One Child



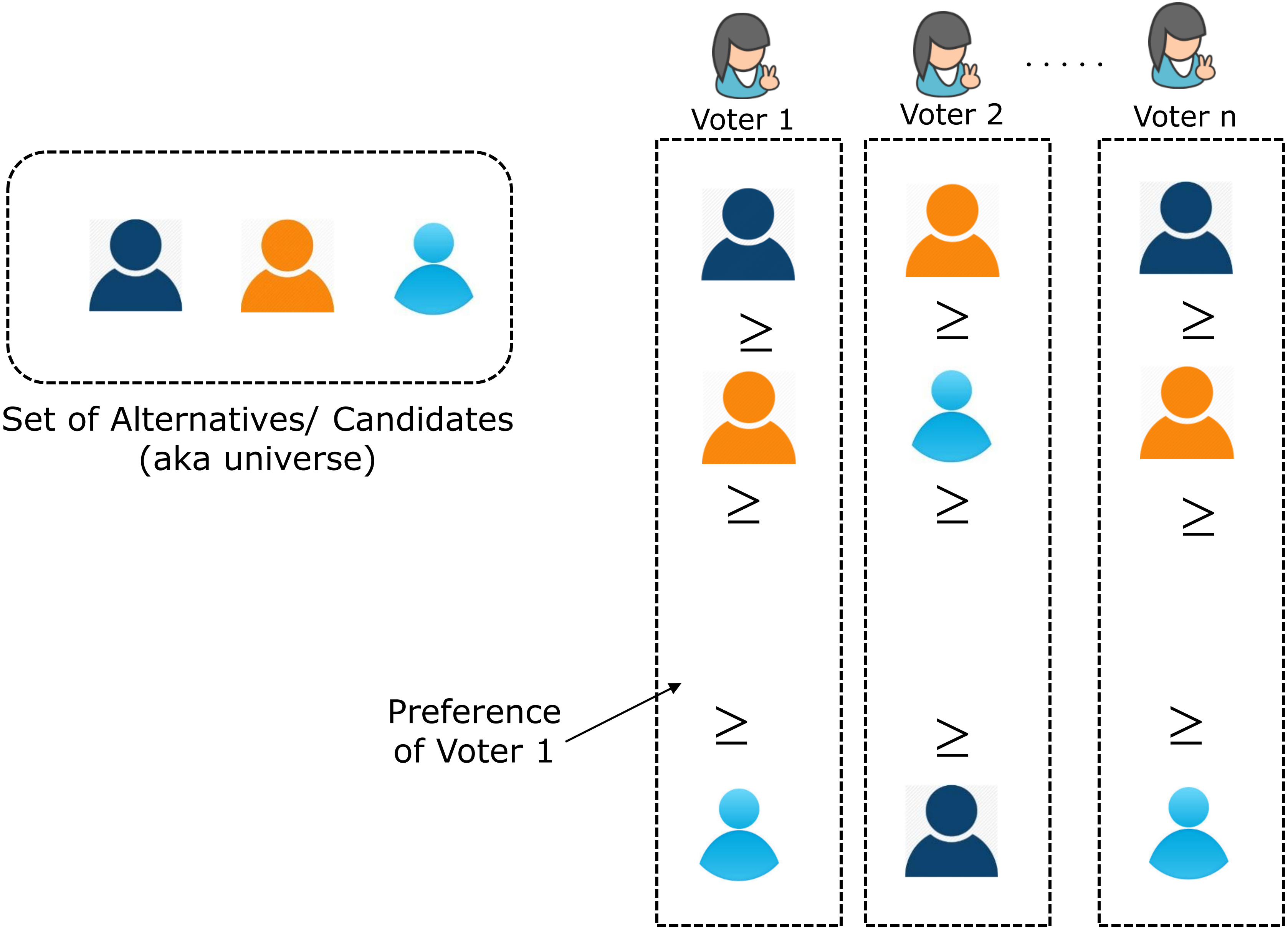
Mechanism Designer
(Birbal/ Tenali Rama/ King Solomon)

Mother 1
(Rational & Intelligent)

Baby

Mother 2
(Rational & Intelligent)

Example 3: Voting



Voter 1

Voter 2

. . . . .

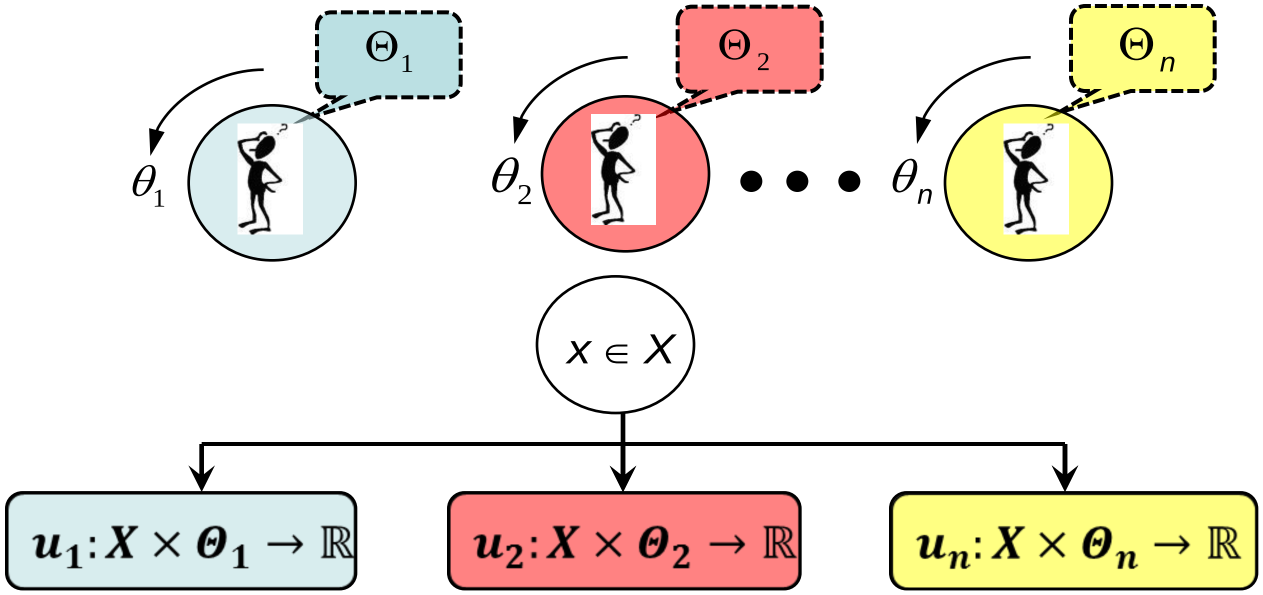
Voter n

Set of Alternatives/ Candidates
(aka universe)

Preference
of Voter 1

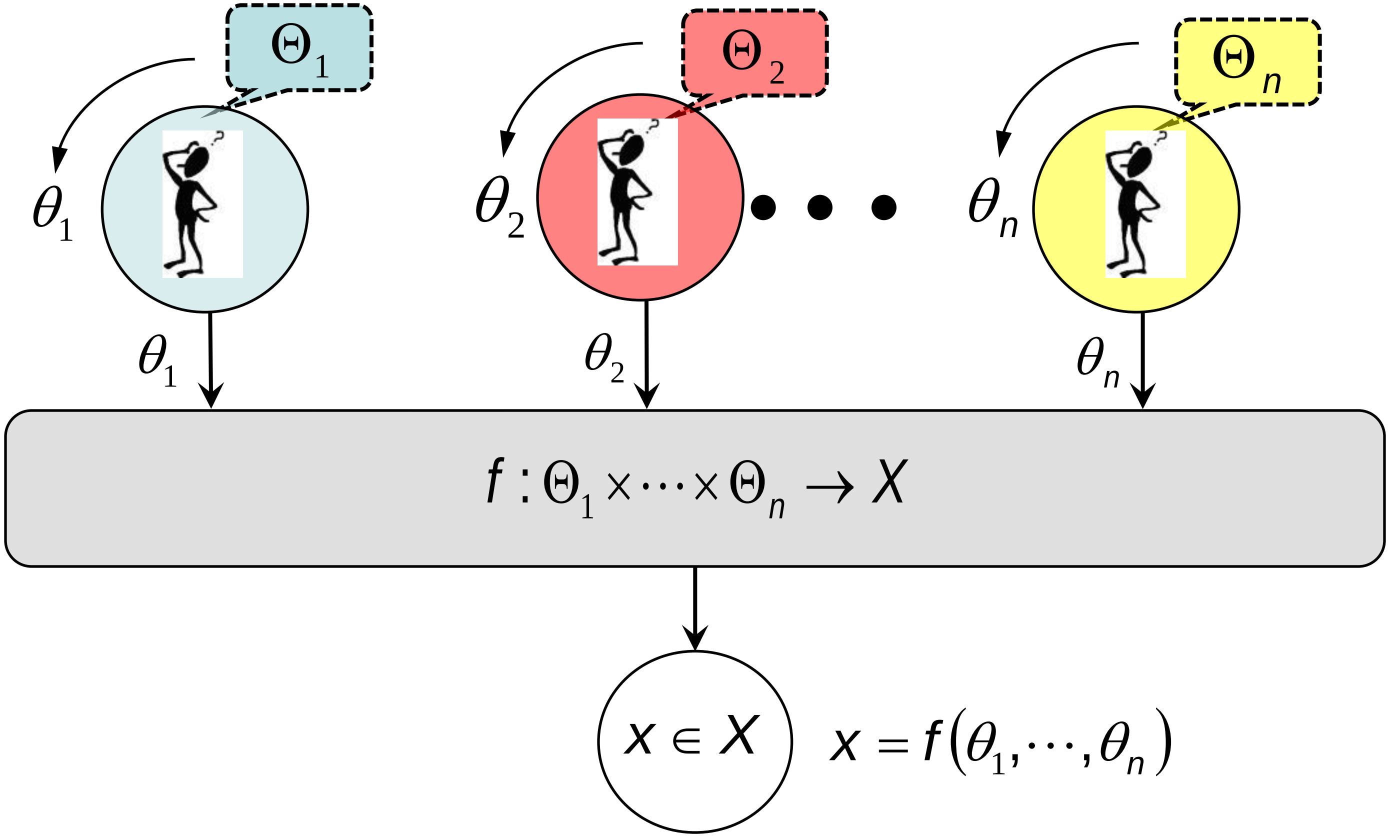
Example 4: Auctions

Mechanism Design Setup



- n agents need to make a collective choice from outcome set X
- Each agent i privately observes a signal θ_i
- Signal θ_i determines agent i's preferences over the set X
- Signal θ_i is known as agent i's type.
- The set of agent i's possible types is denoted by Θ_i
- Agent types, $\theta = (\theta_1, \theta_2, \cdots, \theta_n)$ are drawn according to a distribution $\Phi(.)$
- Each agent is rational, intelligent, and tries to maximize his utility $u_i : X \times \Theta_i \to \mathbb{R}$
- $\Phi(.), \Theta_1, \cdots \Theta_n, u_1(.), \cdots u_n(.)$ are common knowledge among the agents

Social Choice Function (SCF)



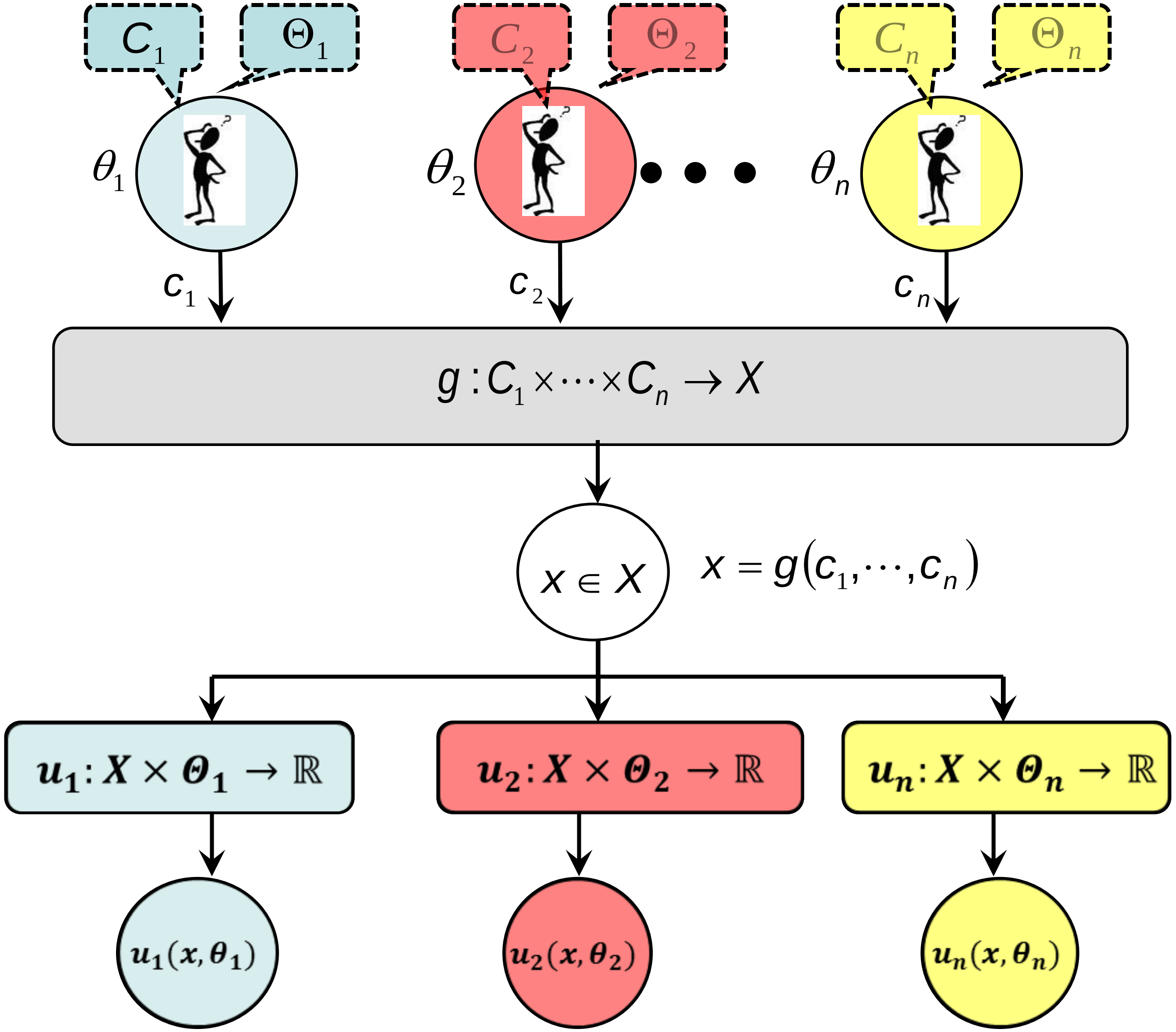
$$f:\Theta_1 \times \cdots \times \Theta_n \rightarrow X$$

$$x = f(\theta_1, \cdots, \theta_n)$$

Planner ideally wants to aggregate preferences as per SCF

(had he known true types of all the agents)

So What is Mechanism ?



C_1 Θ_1 C_2 Θ_2 C_n Θ_n

θ_1 θ_2 $\bullet\bullet\bullet$ θ_n

c_1 c_2 c_n

$g: C_1 \times \cdots \times C_n \to X$

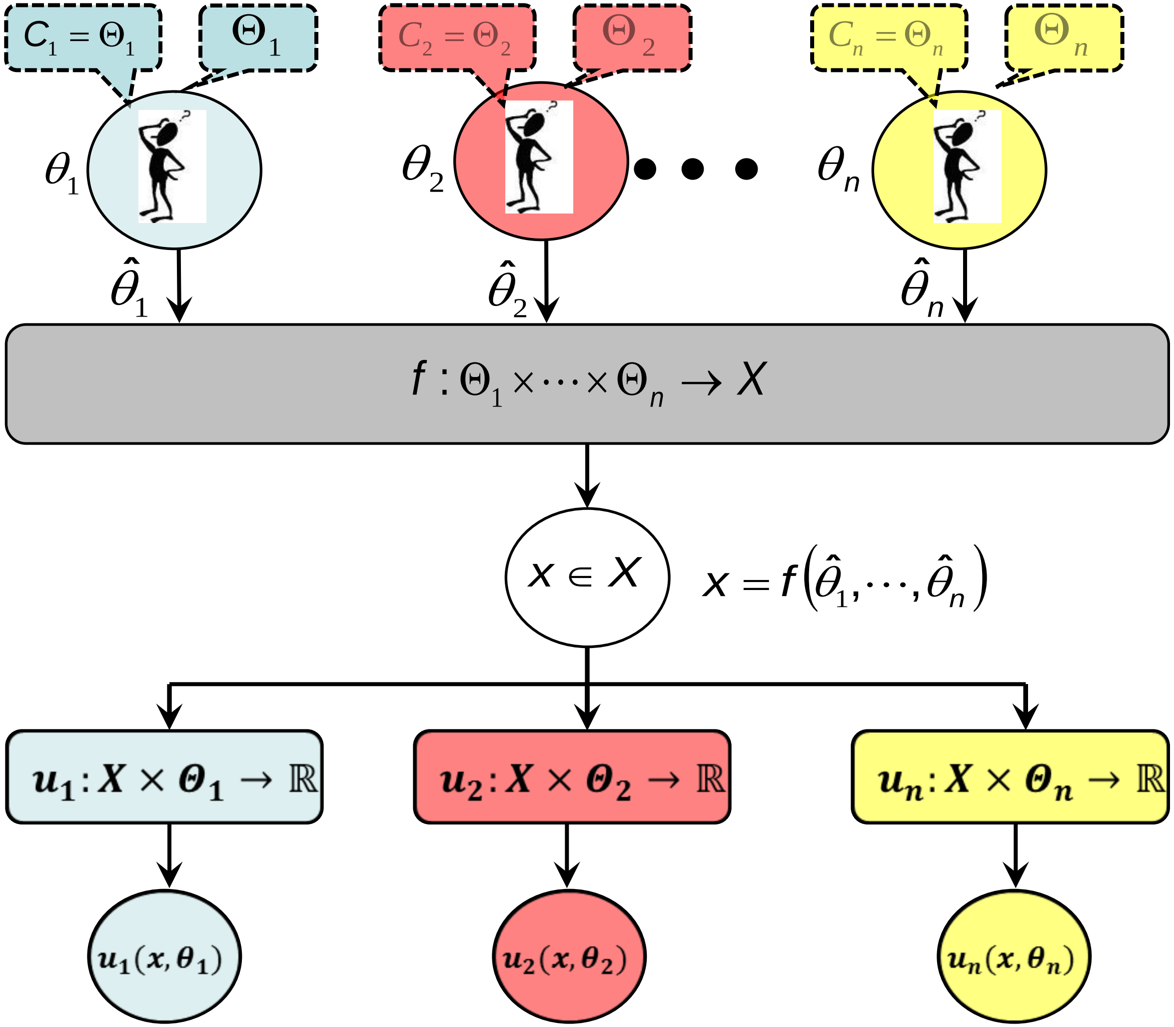
$x \in X$ $x = g(c_1, \cdots, c_n)$

$\boldsymbol{u_1 : X \times \Theta_1} \to \mathbb{R}$ $\boldsymbol{u_2 : X \times \Theta_2} \to \mathbb{R}$ $\boldsymbol{u_n : X \times \Theta_n} \to \mathbb{R}$

$\boldsymbol{u_1(x, \theta_1)}$ $\boldsymbol{u_2(x, \theta_2)}$ $\boldsymbol{u_n(x, \theta_n)}$

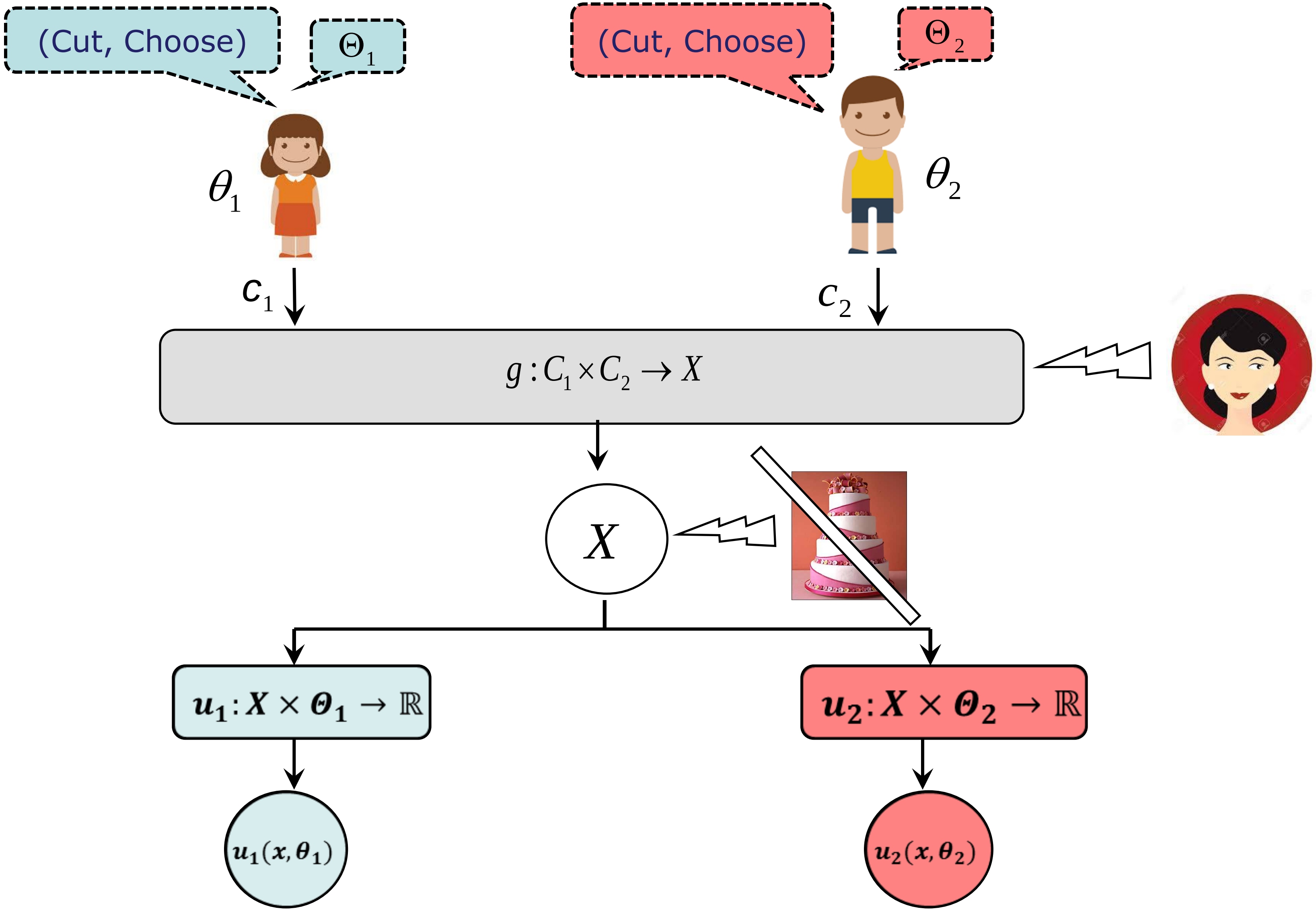
An Indirect Mechanism $M = (g(.), (C_i)_{i \in N})$

Direct Revelation Mechanism (DRM)

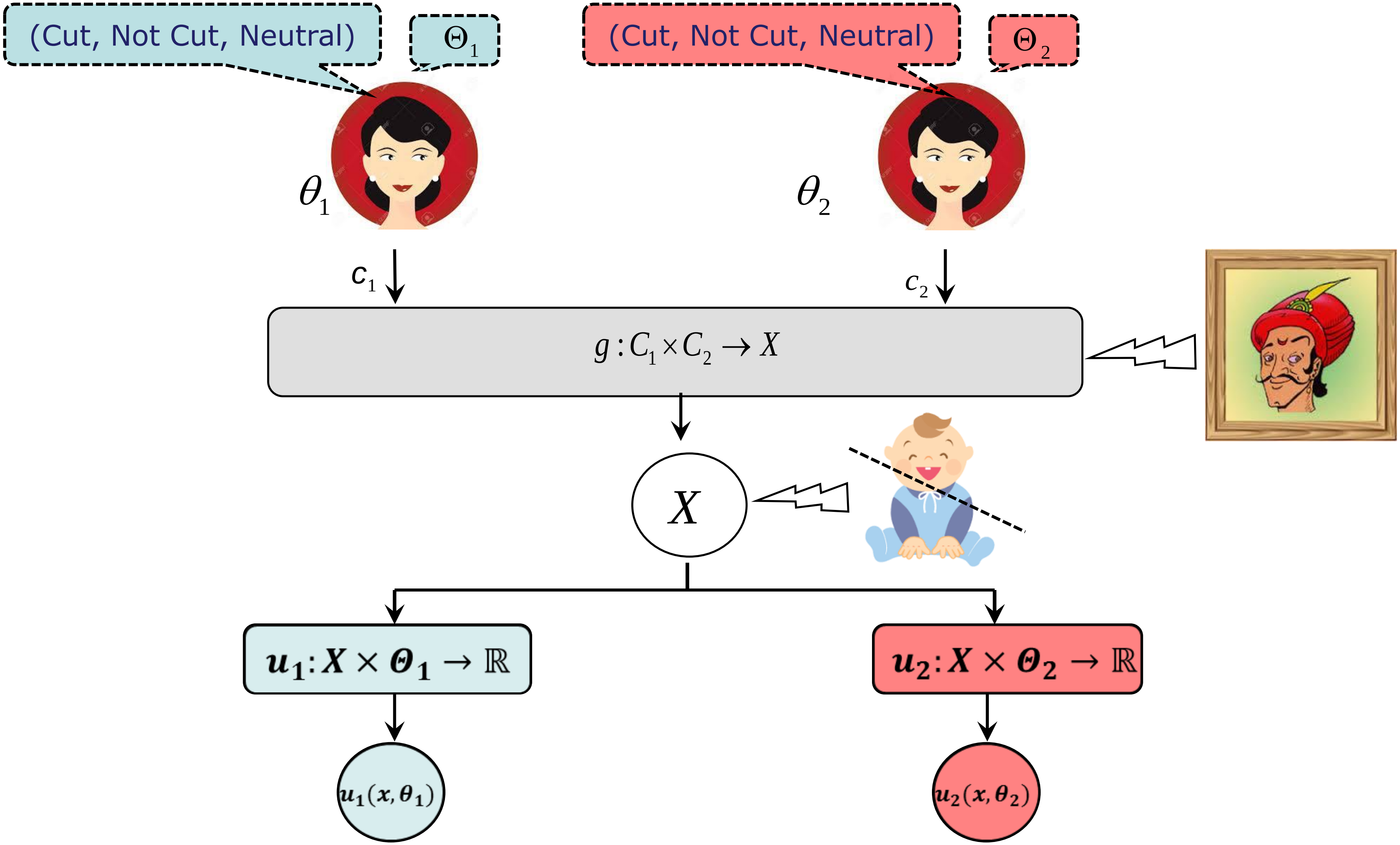


Direct Revelation Mechanism (DRM): $M = (f(.), (\Theta_i)_{i \in N})$

Example 1: Cake Cutting Problem



Example 2: Two Mothers One Child



(Cut, Not Cut, Neutral)

Θ_1

(Cut, Not Cut, Neutral)

Θ_2

θ_1

θ_2

c_1

c_2

$g: C_1 \times C_2 \to X$

X

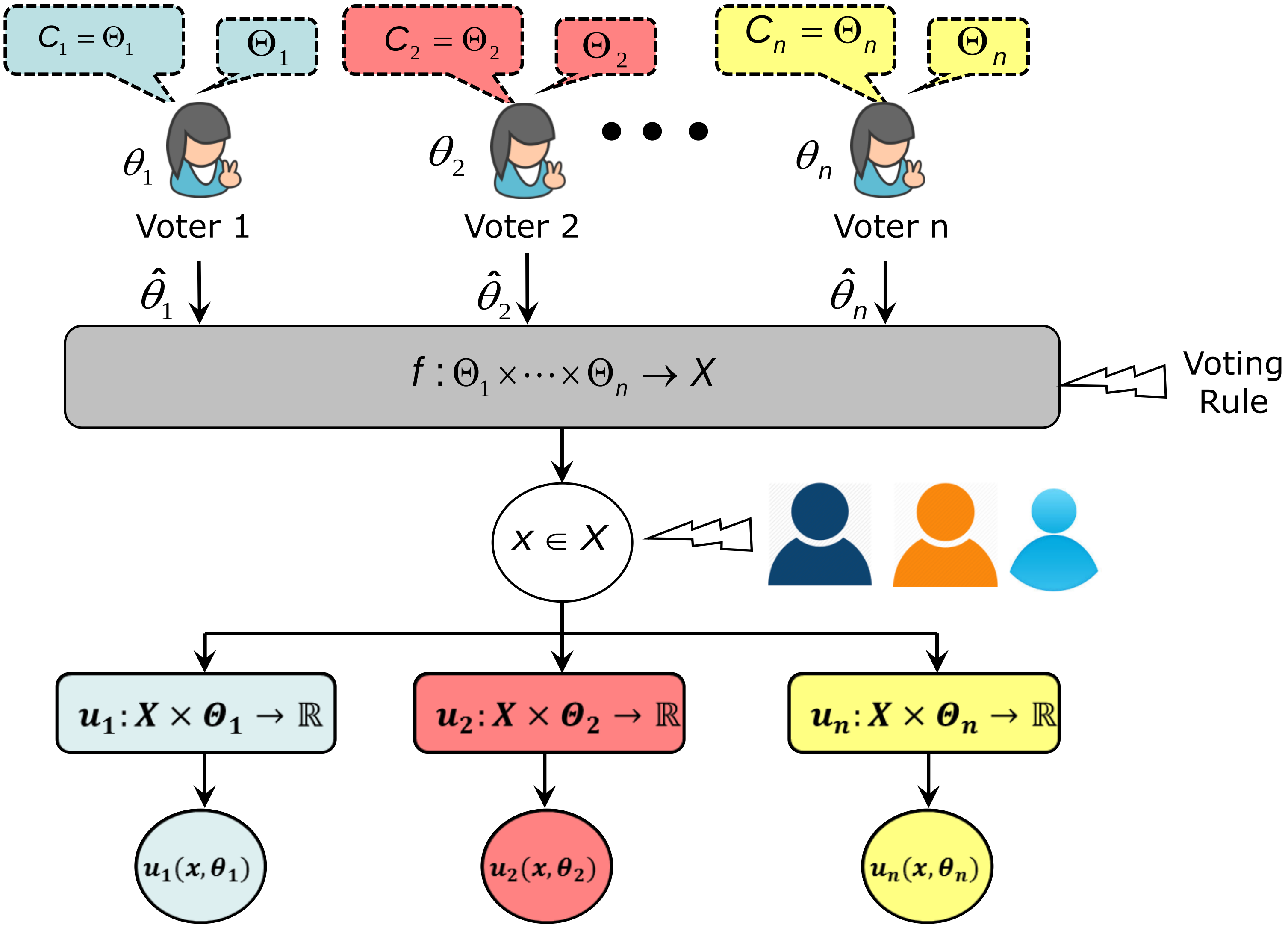
$\boldsymbol{u_1}: \boldsymbol{X} \times \boldsymbol{\Theta_1} \to \mathbb{R}$

$\boldsymbol{u_2}: \boldsymbol{X} \times \boldsymbol{\Theta_2} \to \mathbb{R}$

$\boldsymbol{u_1(x, \theta_1)}$

$\boldsymbol{u_2(x, \theta_2)}$

Example 3: Voting



$C_1 = \Theta_1$ Θ_1 $C_2 = \Theta_2$ Θ_2 $C_n = \Theta_n$ Θ_n

θ_1 θ_2 … θ_n

Voter 1 Voter 2 Voter n

$\hat{\theta}_1$ $\hat{\theta}_2$ $\hat{\theta}_n$

$f : \Theta_1 \times \cdots \times \Theta_n \to X$

Voting Rule

$x \in X$

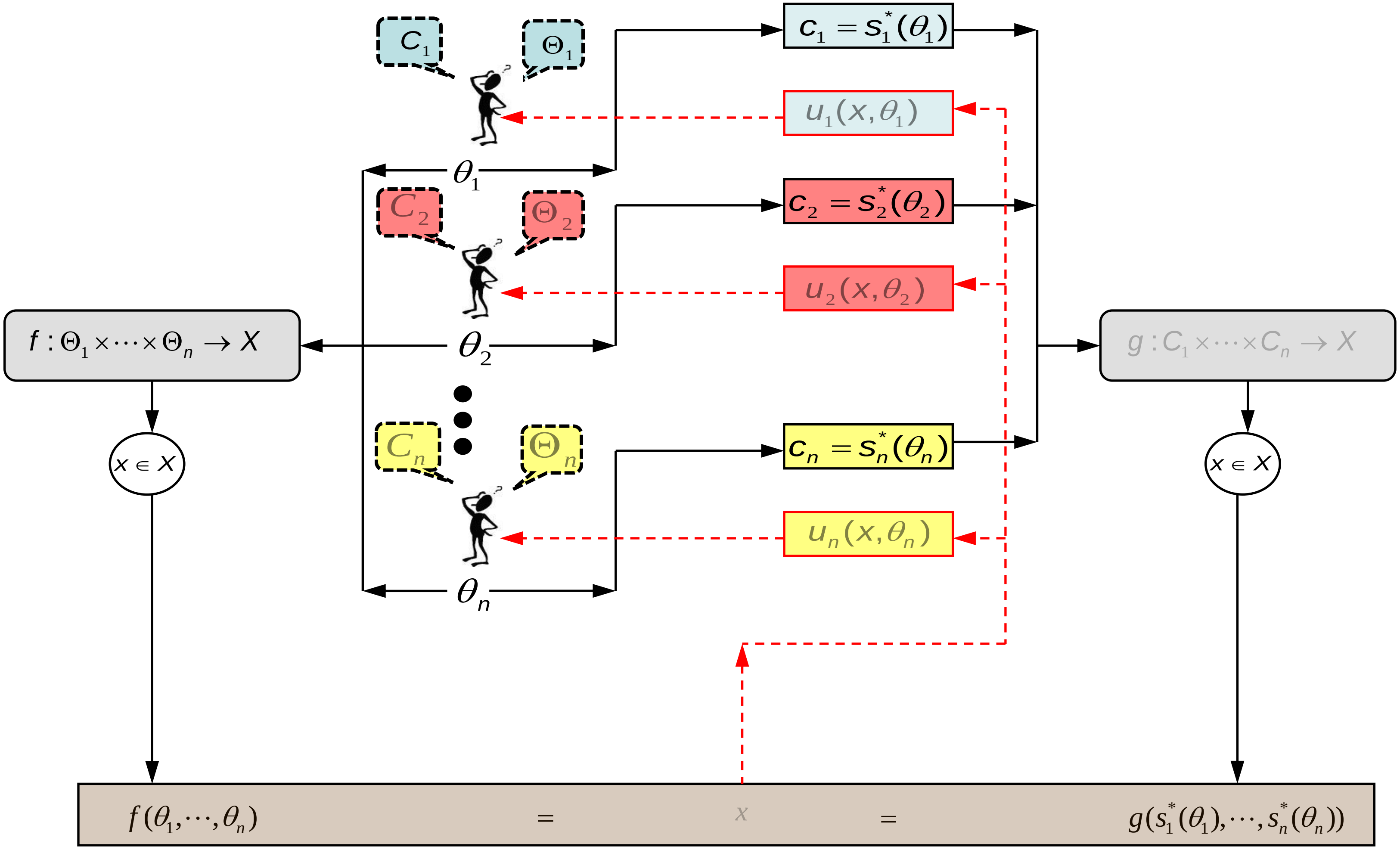
$\boldsymbol{u_1 : X \times \Theta_1} \to \mathbb{R}$ $\boldsymbol{u_2 : X \times \Theta_2} \to \mathbb{R}$ $\boldsymbol{u_n : X \times \Theta_n} \to \mathbb{R}$

$\boldsymbol{u_1(x, \theta_1)}$ $\boldsymbol{u_2(x, \theta_2)}$ $\boldsymbol{u_n(x, \theta_n)}$

Implementing an SCF via Mechanism

Mechanism $M = \left(g(.),\ (C_i)_{i\in N}\right)$

Induced Bayesian Game $\Gamma^b = \left(N, (C_i)_{i\in N}, (\Theta_i)_{i\in N}, \phi(.), (u_i)_{i\in N}\right)$



Equilibrium of Induced Bayesian Game

- **Dominant Strategy Equilibrium (DSE)**

A strategy profile $\left(s_1^d(.), \cdots s_n^d(.)\right)$ is said to be dominant strategy equilibrium if

$$u_i(g(s_i^d(\theta_i), s_{-i}(\theta_{-i})), \theta_i) \geq u_i(g(s_i(\theta_i), s_{-i}(\theta_{-i})), \theta_i)$$

$$\forall i \in N, \theta_i \in \Theta_i, s_i \in S_i, s_{-i} \in S_{-i}$$

- **Bayesian Nash Equilibrium (BNE)**

A strategy profile $\left(s_1^*(.), \cdots s_n^*(.)\right)$ is said to be Bayesian Nash equilibrium

$$E_{\theta_{(-i)}}[u_i(g(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i) \mid \theta_i] \geq E_{\theta_{(-i)}}[u_i(g(s_i(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i) \mid \theta_i]$$

$$\forall i \in N, \theta_i \in \Theta_i, s_i \in S_i$$

- **Observation**

Dominant Strategy-equilibrium $\Longrightarrow$ Bayesian Nash- equilibrium

Implementing an SCF

- **Dominant Strategy Implementation**

We say that mechanism $M = \left(g(.), (C_i)_{i \in N}\right)$ implements SCF $f : \Theta \rightarrow X$ in dominant strategy equilibrium if

$$g\left(s_1^d(\theta_1), \cdots s_n^d(\theta_n)\right) = f(\theta_1, \cdots, \theta_n) \qquad \forall (\theta_1, \cdots, \theta_n)$$

- **Bayesian Nash Implementation**

We say that mechanism $M = \left(g(.), (C_i)_{i \in N}\right)$ implements SCF $f : \Theta \rightarrow X$ in Bayesian Nash equilibrium if

$$g\left(s_1^*(\theta_1), \cdots s_n^*(\theta_n)\right) = f(\theta_1, \cdots, \theta_n) \qquad \forall (\theta_1, \cdots, \theta_n)$$

- **Observation**

Dominant Strategy-implementation $\Longrightarrow$ Bayesian Nash- implementation

---

- *Andreu Mas Colell, Michael D. Whinston, and Jerry R. Green, “**Microeconomic Theory**”, Oxford University Press, New York, 1995.*

Properties of an SCF

- **(Ex Post) Efficiency**

  For no profile of agents' type $\theta = (\theta_1, \cdots, \theta_n)$ there exists an $x \in X$ such that $u_i(x, \theta_i) \geq u_i(f(\theta), \theta_i) \ \forall\, i$ and $u_i(x, \theta_i) > u_i(f(\theta), \theta_i)$ for some i

  (Deviation from SCF recommended outcome can't make someone better-off without making anyone else worse-off)

- **Dominant Strategy Incentive Compatibility (DSIC)**

  If the direct revelation mechanism $D = (f(.), (\Theta_i)_{i \in N})$ has a dominant strategy equilibrium $(s_1^d(.), \cdots s_n^d(.))$ in which

  $$s_i^d(\theta_i) = \theta_i, \forall\, \theta_i \in \Theta_i, i \in N$$

  (Irrespective of what others are doing, I must admit trust)

- **Bayesian Incentive Compatibility (BIC)**

  If the direct revelation mechanism $D = (f(.), (\Theta_i)_{i \in N})$ has a Bayesian Nash equilibrium $(s_1^*(.), \cdots s_n^*(.))$ in which

  $$s_i^*(\theta_i) = \theta_i, \forall\, \theta_i \in \Theta_i, i \in N$$

  (If others are admitting truth, I must also do the same)

Properties of an SCF

- **Dictatorial**

For every profile of agents' type $\theta = (\theta_1, \cdots, \theta_n)$, we have

$$f(\theta) = \{x \in X \mid u_d(x, \theta_d) \geq u_d(y, \theta_d) \forall y \in X\}$$

where d is a particular agent known as dictator.

(A special agent is favored all the times by the planner)

- **(Ex Post) Individual Rationality**

$$u_i(f(\theta_i, \theta_{-i}), \theta_i) \geq \overline{u_i}(\theta_i) \quad \forall(\theta_i, \theta_{-i})$$

where $\overline{u_i}(\theta_i)$ is the utility that agent i receives by withdrawing from the mechanism when his type is θ_i

(Participation in the mechanism will never make anyone worse-off)

Gibbard – Satterthwaite Impossibility Theorem

GS Theorem: Suppose for a given SCF

(1) Range is finite and contains at least 3 elements
(2) Preference structure (aka type space) is rich

then, the SCF is DSIC iff it is dictatorial

(By and large, DSIC and Non-dictatorship don't co-exist)

Possible Ways Out

1. Relax the assumption on richness of preferences (e.g. single peaked preferences)
2. Relax the assumption on finite range by allowing transfer of payments
3. Relax the requirement of strong solution concept namely DSIC and instead work with BIC

- A. Gibbard. Manipulation of voting schemes. *Econometrica*, 41:587-601, 1973.
- M. A. Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorem for voting procedure and social welfare functions. *Journal of Economic Theory*, 10:187-217, 1975.

Quasi-Linear Environment

$$X = \left\{ (k, t_1, \cdots, t_n) \mid k \in K, t_i \in \Re \ \forall i = 1, \cdots, n, \sum_i t_i \le 0 \right\}$$

project choice

Monetary transfer
to agent i

No outside source of funding

$$u_1(x, \theta_1) = v_1(k, \theta_1) + t_1$$

Valuation function of agent 1

Properties of an SCF in Quasi-Linear Environment

- **Allocative Efficiency (AE)**

An SCF $f(.) = (k(.), t_1(.), \cdots, t_n(.))$ is AE if for each $\theta \in \Theta$, $k(\theta)$ satisfies

$$k(\theta) \in \arg\max_{k \in K} \sum_{i=1}^{n} v_i(k, \theta_i)$$

- **(Strong) Budget Balance**

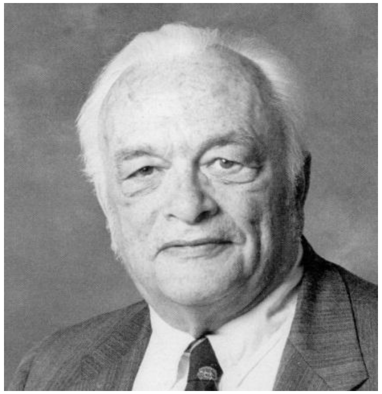
An SCF $f(.) = (k(.), t_1(.), \cdots, t_n(.))$ is BB if for each $\theta \in \Theta$, we have

$$\sum_{i=1}^{n} t_i(\theta) = 0$$

Lemma

An SCF $f(.) = (k(.), t_1(.), \cdots, t_n(.))$ is (ex post) efficient in quasi-linear environment (having no outside source of funding) iff it is (AE + BB)

Vickrey-Clarke-Groves (VCG) Mechanisms



Vickrey (1961)

Clarke (1971)

Groves (1973)

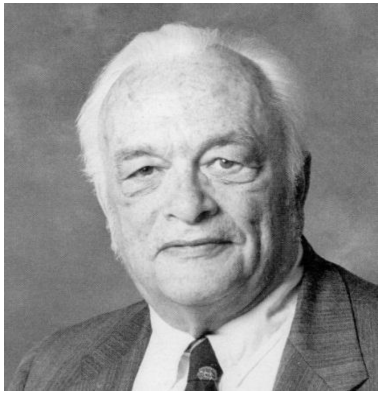
Groves Theorem [1973]: Let the SCF $f(.) = \left(k^*(.), t_1(.), \cdots, t_n(.)\right)$ be AE. This SCF can be truthfully implemented in dominant strategy if it satisfied the following payment structure

$$t_i(\theta) = \left[\sum_{j\neq i} v_j(k^*(\theta), \theta_j)\right] + h_i(\theta_{-i}) \forall i = 1, \cdots n$$

Clarke Pivotal Mechanism [1973]: $h_i(\theta_{-i}) = \left[\sum_{j\neq i} v_j(k^*_{-i}(\theta), \theta_j)\right]$

- T. Groves. Incentives in teams. *Econometrica*, 41:617-631, 1973.
- E. Clarke. Multi-part pricing of public goods. *Public Choice*, 11:17-23, 1971.
- W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8-37, March 1961.

Vickrey-Clarke-Groves (VCG) Mechanisms



Vickrey

Clarke

Groves

Under some mild conditions on preference structure, VCG are the only mechanisms in quasi-linear environment satisfying AE+DSIC

(Im) Possibility Theorems in Quasi-Linear Environments

Groves (Possibility) Theorem

In any quasi-linear environment, there exists an SCF which is *AE + DSIC*

Green- Laffont (Impossibility) Theorem

In any quasi-linear environment, if preference structure (aka type space) is rich then there is no SCF which is *AE + BB+ DSIC*

The dAGVA (Possibility) Theorem

In any quasi-linear environment, there exists a social choice function which is *AE+BB+BIC*

(Im) Possibility Theorems in Quasi-Linear Environments

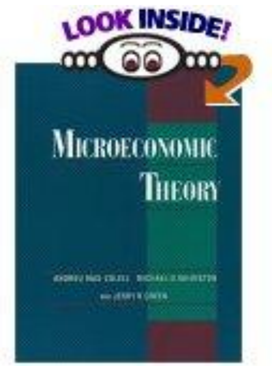
Myerson-Satterthwaite (Impossibility) Theorem
In the quasi-linear environment, there is no SCF which is AE + BB+ BIC +IR

Myerson's (Possibility) Theorem for Optimal Mechanism
In the quasi-linear environment, if the type is one dimensional, then there exist SCF which are BIC+ IIR and maximize the earning (or surplus) of the planner

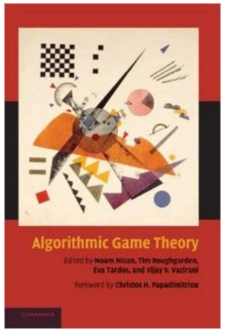
---

- *R. B. Myerson. Optimal Auction Design. Math. Operations Res., 6(1): 58 -73, Feb. 1981.*

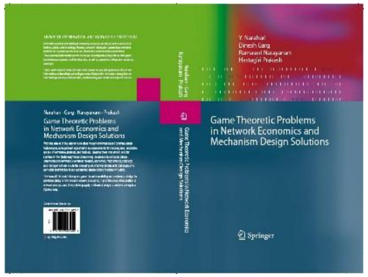
References



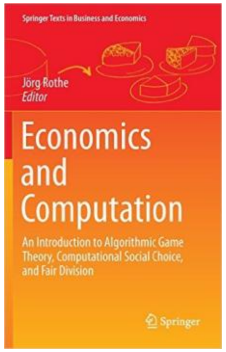
Andrew Mas-Collel, Michael D. Whinston, and Jerry R. Green (1995), “***Microeconomic Theory***”, Oxford University Press



Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay Vazirani (2007), “***Algorithmic Game Theory***”, Cambridge University Press



Y. Narahari, Dinesh Garg, Ramasuri Narayanam, and Hastagiri Prakash (2009), “***Game Theoretic Problems in Network Economics and Mechanism Design***”, Springer



Jorg Rothe (2015), “Economics and Computation: ***An Introduction to Algoritmic Game Theory, Computational Social Choice, and Fair Division***”, Springer

Thank you

Two Fundamental Design Aspects

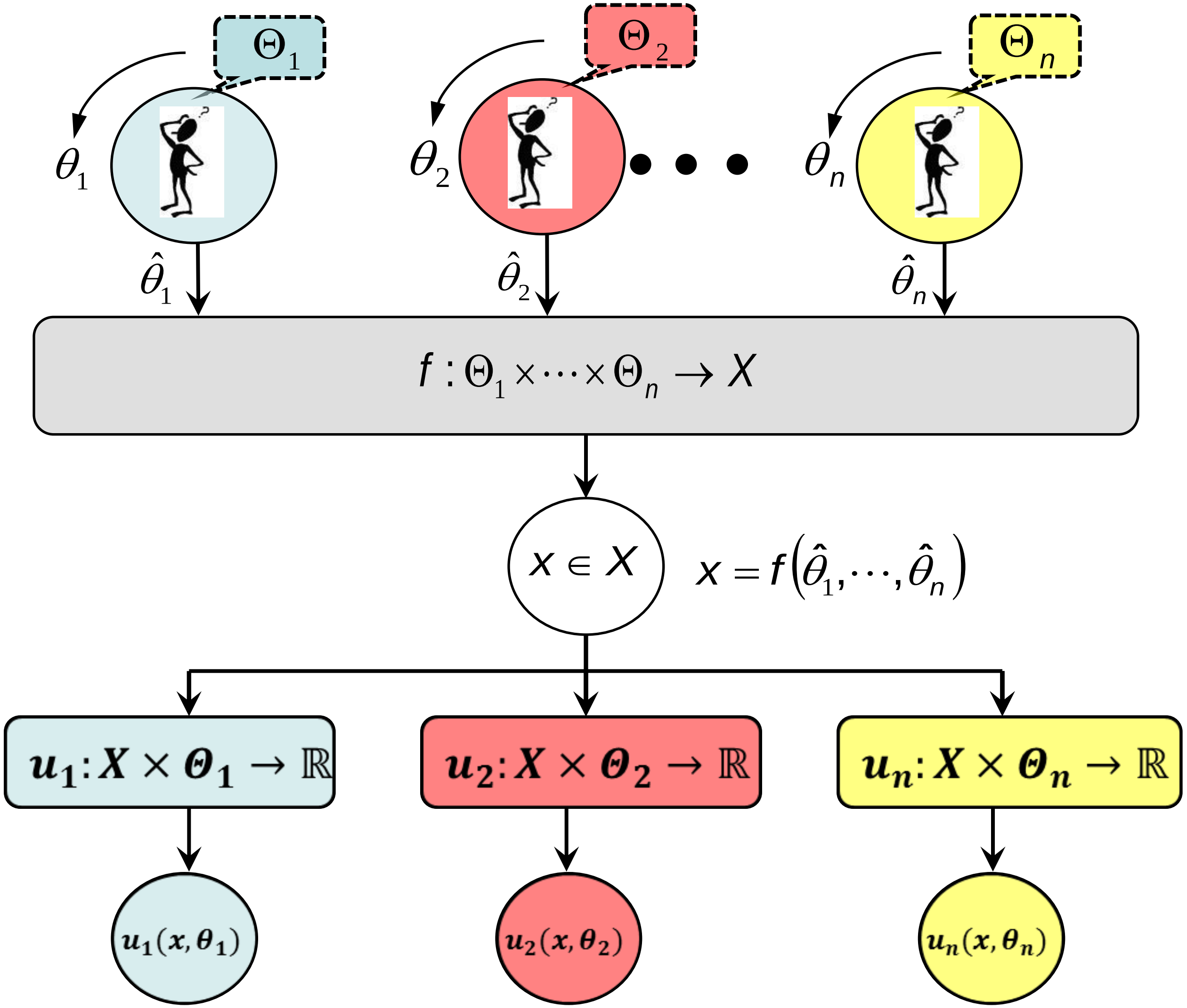
- **Preference Aggregation**

  For a given type profile $\boldsymbol{\theta} = (\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \cdots, \boldsymbol{\theta_n})$ of the agents, what outcome $x \in \boldsymbol{X}$ should be chosen ?

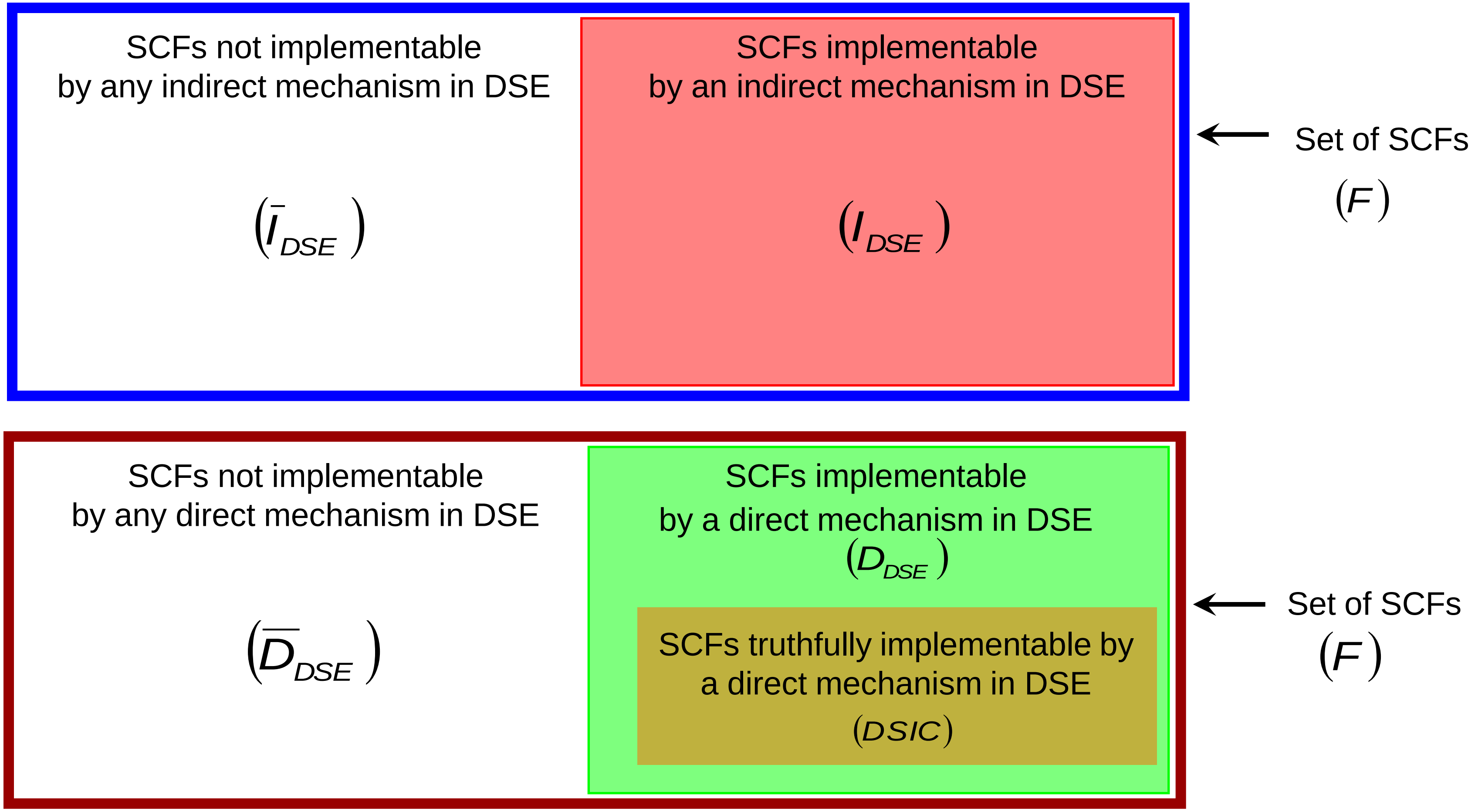
- **Information Revelation (Elicitation)**

  How do we elicit the true type $\boldsymbol{\theta_i}$ of each agent $\boldsymbol{i}$, which is his private information ?

Information Elicitation Problem



Revelation Principle for DSE



SCFs not implementable
by any indirect mechanism in DSE

$(\bar{I}_{DSE})$

SCFs implementable
by an indirect mechanism in DSE

(I_{DSE})

← Set of SCFs

(F)

SCFs not implementable
by any direct mechanism in DSE

$(\overline{D}_{DSE})$

SCFs implementable
by a direct mechanism in DSE
(D_{DSE})

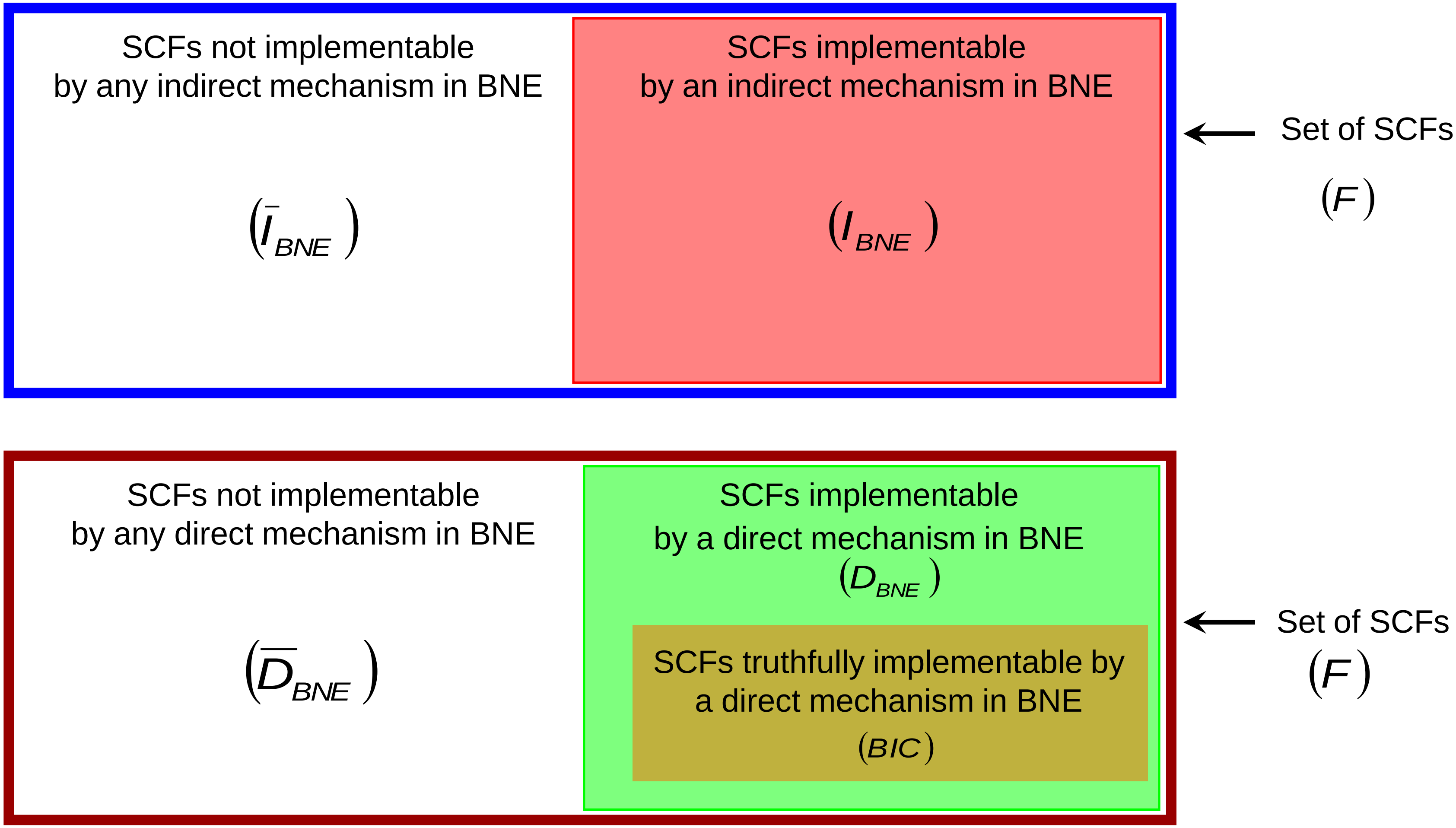
SCFs truthfully implementable by
a direct mechanism in DSE
$(DSIC)$

← Set of SCFs

(F)

$$DSIC \subseteq D_{DSE} \subseteq \boxed{I_{DSE} \subseteq DSIC}$$

Revelation Principle for BNE



SCFs not implementable
by any indirect mechanism in BNE

$(\bar{I}_{BNE})$

SCFs implementable
by an indirect mechanism in BNE

(I_{BNE})

← Set of SCFs

(F)

SCFs not implementable
by any direct mechanism in BNE

$(\bar{D}_{BNE})$

SCFs implementable
by a direct mechanism in BNE
(D_{BNE})

SCFs truthfully implementable by
a direct mechanism in BNE
(BIC)

← Set of SCFs
(F)

$$BIC \subseteq D_{BNE} \subseteq I_{BNE} \subseteq BIC$$

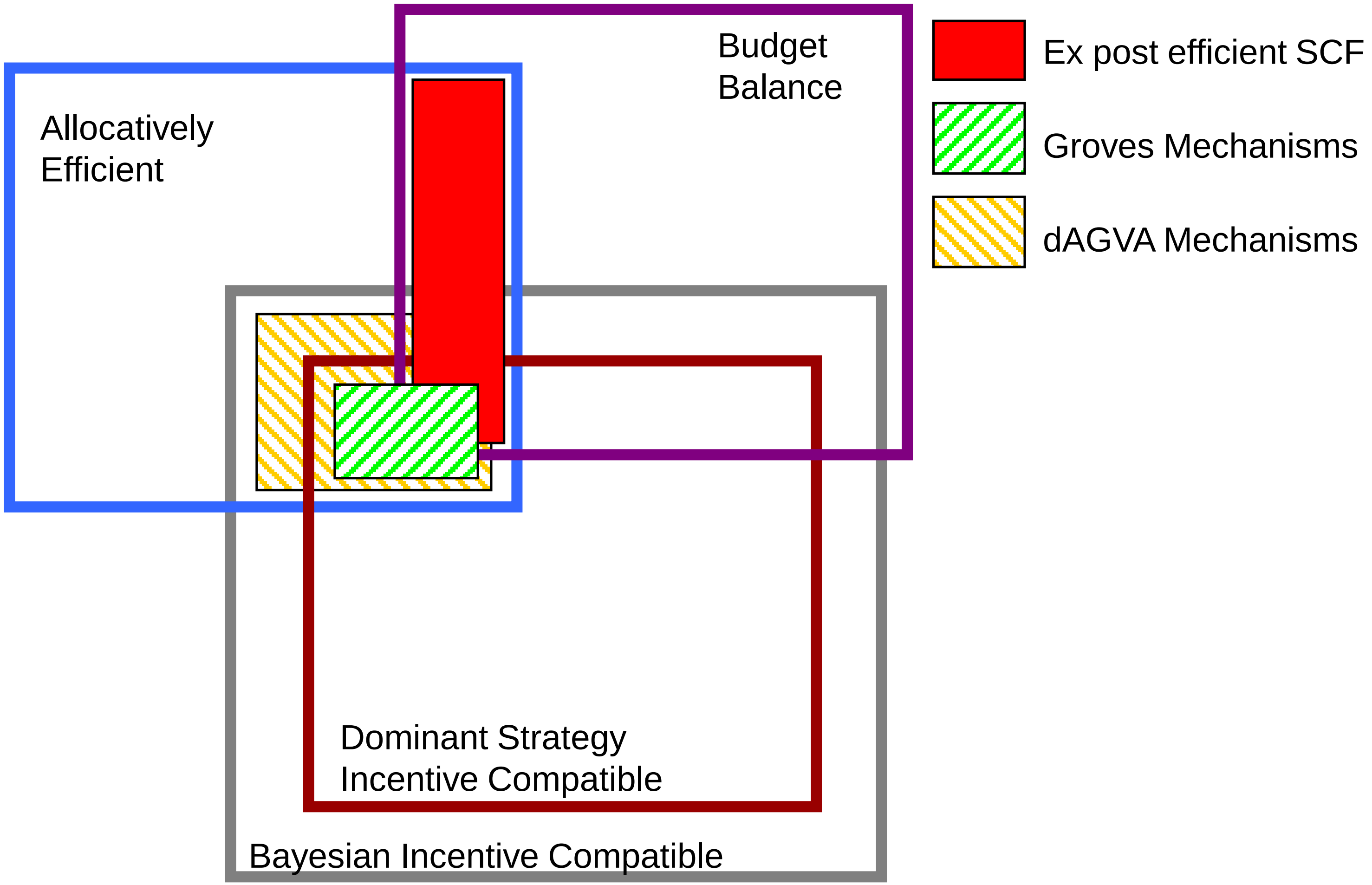
Absence of Dictatorial SCF in Quasi-Linear Environments

Lemma:

In quasi-linear environment (having no source of outside funding), the utility of an agent can be made arbitrary high and thereby no SCF is a dictatorial SCF in this environment

(Thus, GS impossibility theorem does not bite us here)

Space of SCFs in Quasi-linear Environment



Budget Balance

Allocatively Efficient

Dominant Strategy Incentive Compatible

Bayesian Incentive Compatible

🟥 Ex post efficient SCF

▨ Groves Mechanisms

▨ dAGVA Mechanisms