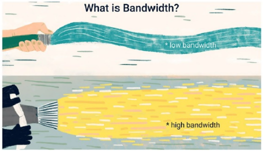
Game theory for cumulative-prospect-theoretic agents

Venkat Anantharam
based on the thesis work of Soham Phade
University of California, Berkeley

Workshop on Learning in Games
Institut de Mathématiques de Toulouse
Toulouse, France
July 1-3, 2024

Soham Phade

Network economics today

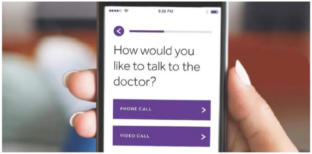


Broadband network

Cloud Computing

Smart Grid

Ad Auctions



Telemedicine

Labor Markets

How to incorporate human preferences?

Decision Making under Uncertainty

Decision Maker = Choose between different **Lotteries**

$L_1 =$

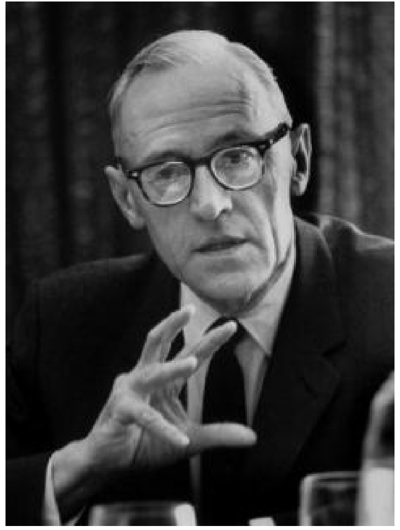
probability	0.1	0.2	0.15	0.1	0.25	0.2
outcome	10	5	2	0	-1	-3

$L_2 =$

probability	0.15	0.05	0.6	0.2
outcome	8	5	-1	-2

Expected Utility Theory (EUT)

(Von Neumann-Morgenstern 1947)



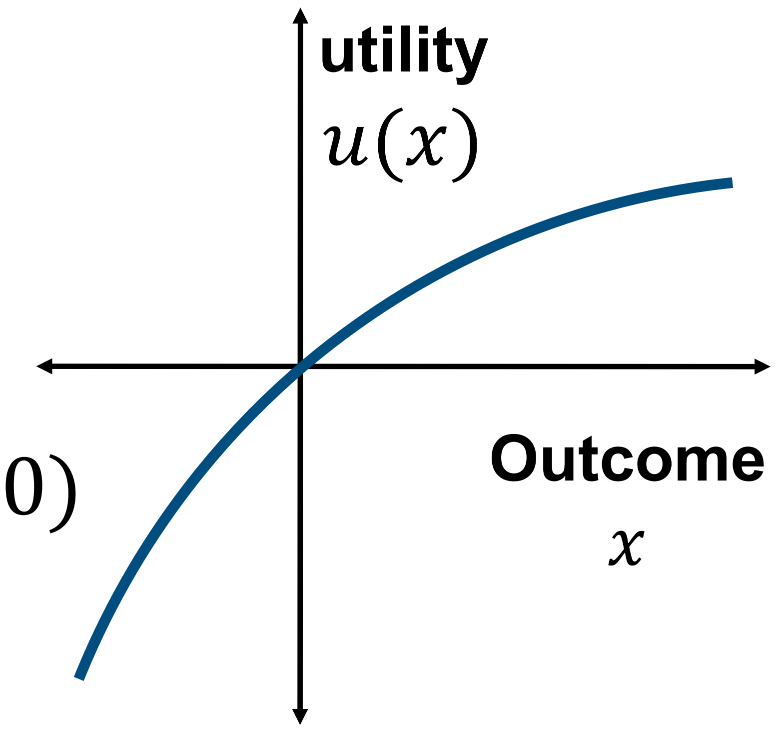
Lottery

$L =$

probability	0.1	0.2	0.15	0.1	0.25	0.2
outcome	10	5	2	0	-1	-3

Expected utility of lottery L

$$U(L) = 0.1u(10) + 0.2u(5) + 0.15u(2) + 0.1u(0) + 0.25u(-1) + 0.2u(-3)$$



Lottery with higher Expected Utility is preferred.

Utility function

Allais Paradox (1953)

Allais

Allais Paradox (1953)

Allais

Experiment 1			
Gamble 1A		Gamble 1B	
Winnings	Chance	Winnings	Chance
$1 million	100%	$1 million	89%
		Nothing	1%
		$5 million	10%

Allais Paradox (1953)

Allais

Experiment 1			
Gamble 1A		Gamble 1B	
Winnings	Chance	Winnings	Chance
$1 million	100%	$1 million	89%
		Nothing	1%
		$5 million	10%

Experiment 2			
Gamble 2A		Gamble 2B	
Winnings	Chance	Winnings	Chance
Nothing	89%	Nothing	90%
$1 million	11%		
		$5 million	10%

Allais Paradox (1953)

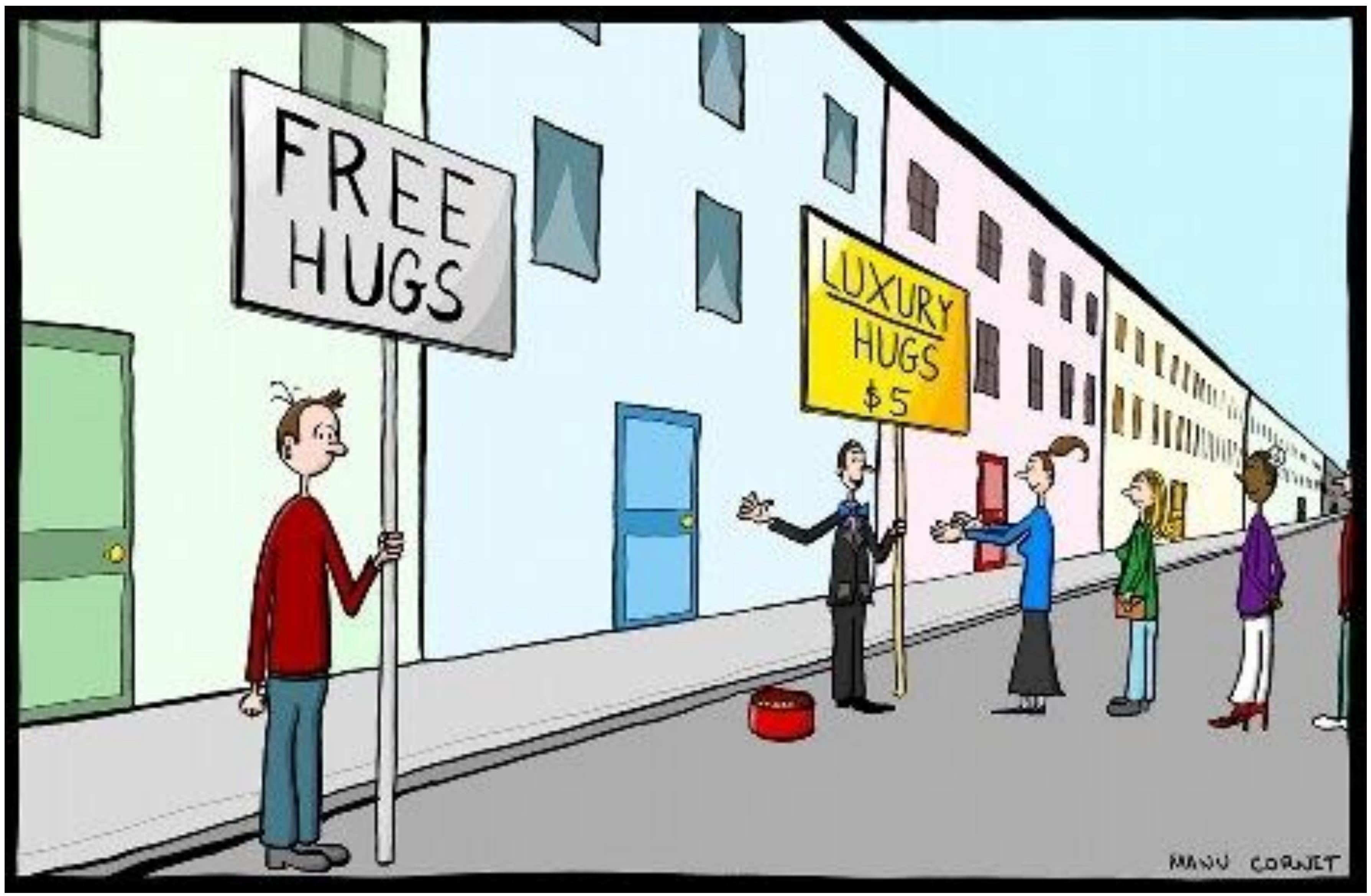
Allais

Experiment 1			
Gamble 1A		Gamble 1B	
Winnings	Chance	Winnings	Chance
$1 million	100%	$1 million	89%
		Nothing	1%
		$5 million	10%

Experiment 2			
Gamble 2A		Gamble 2B	
Winnings	Chance	Winnings	Chance
Nothing	89%	Nothing	90%
$1 million	11%		
		$5 million	10%

People often do NOT follow EUT!

Behavioral Aspects



Cumulative Prospect Theory (CPT)

Kahneman Tversky

Kahneman - Tversky Cumulative Prospect Theory (1992)

Reference Point

$$r \in \mathbb{R}$$

Gains : Outcomes $\geq r$

Losses: Outcomes $< r$

Cumulative Prospect Theory (CPT)


Kahneman Tversky

Kahneman - Tversky Cumulative Prospect Theory (1992)

Reference Point

$$r \in \mathbb{R}$$

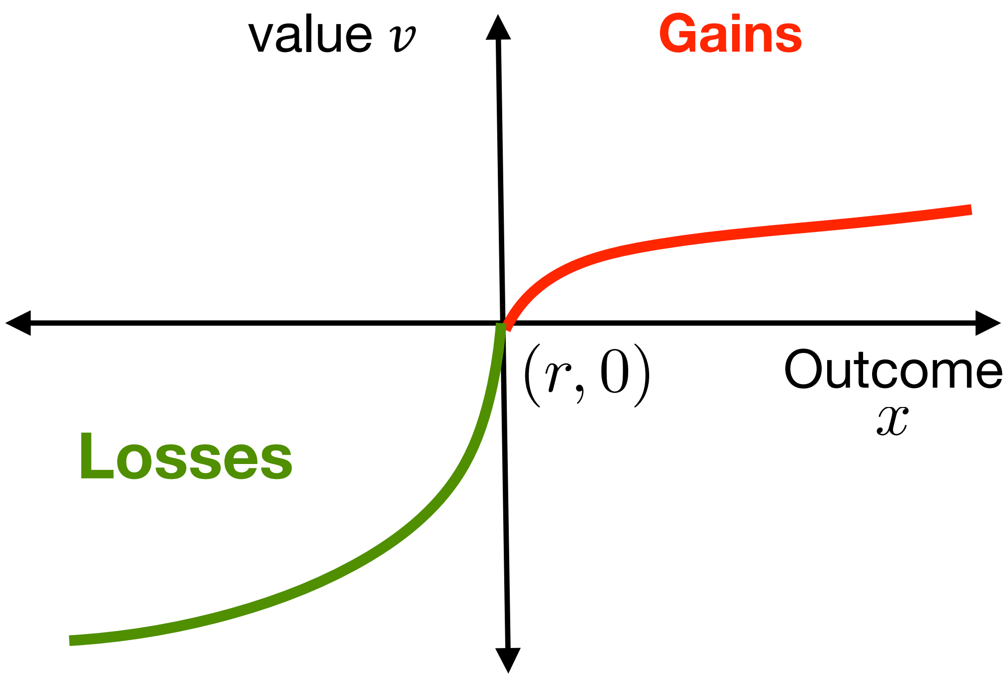
Gains : Outcomes $\geq r$

Losses: Outcomes $< r$

Value Function

$$v : \mathbb{R} \to \mathbb{R}$$

1. $v(x)$ is continuous in x
2. $v(r) = 0$
3. it is strictly increasing in x



Cumulative Prospect Theory (CPT)

Kahneman Tversky

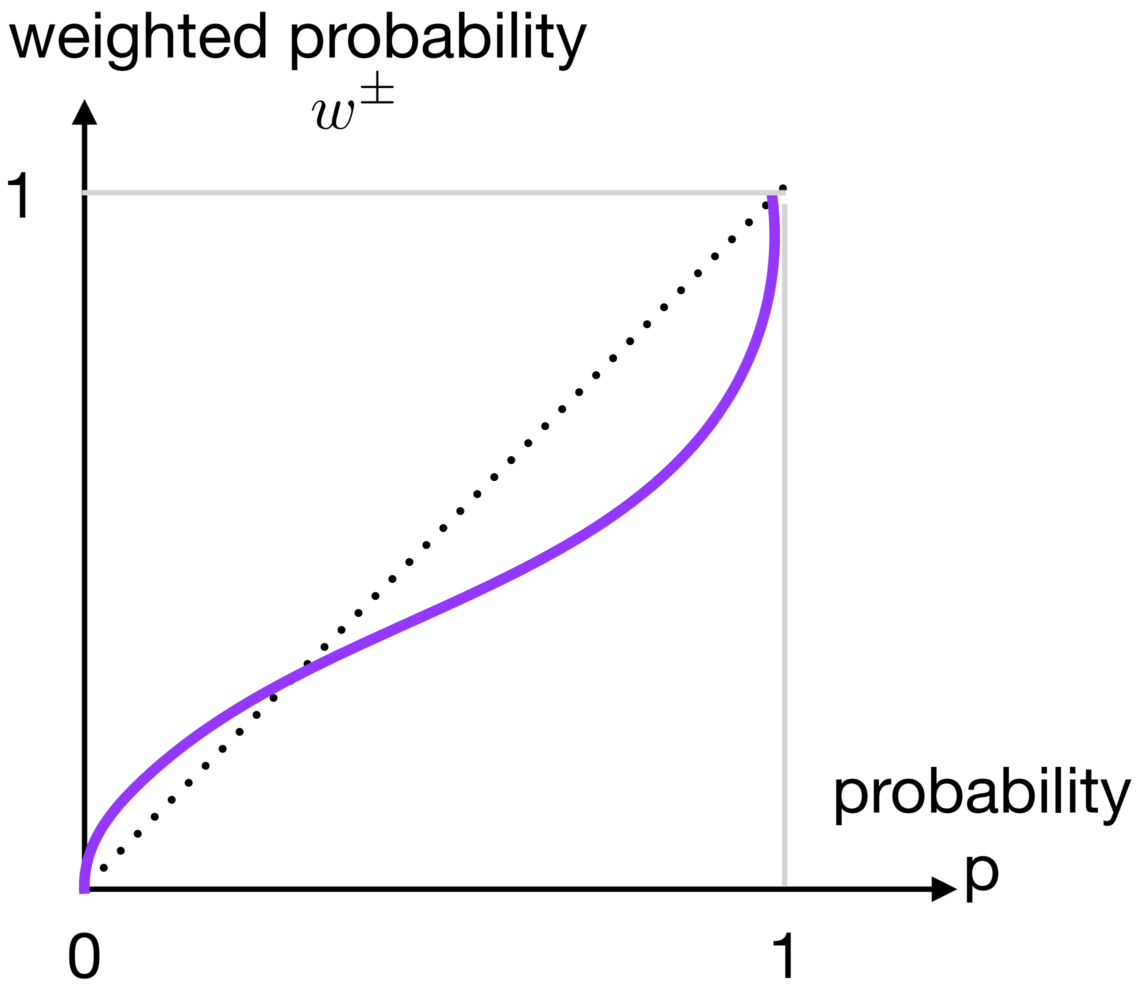
Kahneman - Tversky Cumulative Prospect Theory (1992)

Probability Weighting Functions

$w^+ : [0,1] \to [0,1]$ $w^- : [0,1] \to [0,1]$

Gains Losses

1. they are continuous
2. they are strictly increasing
3. $w^\pm(0) = 0$ and $w^\pm(1) = 1$



Cumulative Prospect Theory (CPT)

Suppose $r = 0$

$L =$

	Gains				Losses	
probability	0.1	0.2	0.15	0.1	0.25	0.2
outcome	10	5	2	0	-1	-3

CPT Value of Lottery L

$$V(L) = V^{gain}(L) + V^{loss}(L)$$

$$V^{gain}(L) = v(10)[w^{+}(0.1)] + v(5)[w^{+}(0.1+0.2) - w^{+}(0.1)] + v(2)[w^{+}(0.1+0.2+0.15) - w^{+}(0.1+0.2)]$$

$$V^{loss}(L) = v(-3)[w^{-}(0.2)] + v(-1)[w^{-}(0.2+0.25) - w^{-}(0.2)]$$

Allais Paradox Resolved

Experiment 1			
Gamble 1A		Gamble 1B	
Winnings	Chance	Winnings	Chance
$1 million	100%	$1 million	89%
		Nothing	1%
		$5 million	10%

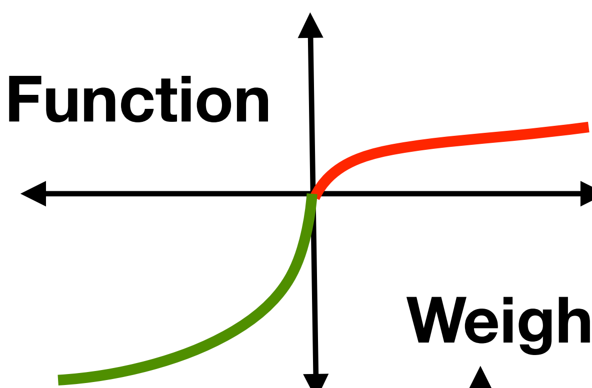
Experiment 2			
Gamble 2A		Gamble 2B	
Winnings	Chance	Winnings	Chance
Nothing	89%	Nothing	90%
$1 million	11%		
		$5 million	10%

Suppose r = $1 million

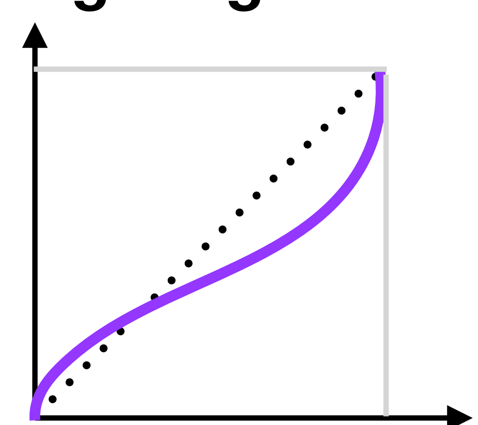
$\mathcal{V}(\text{Lottery 1A}) = 0$
$\mathcal{V}(\text{Lottery 1B}) = v(4)w^{+}(0.10) + v(-1)w^{-}(0.01)$

$\mathcal{V}(\text{Lottery 2A}) = v(-1)w^{-}(0.89)$
$\mathcal{V}(\text{Lottery 2B}) = v(4)w^{+}(0.1) + v(-1)w^{-}(0.90)$

Value Function



Weighting Function



Why CPT?

- Accommodates several empirically observed behavioral features
- Mathematically tractable
- Generalization of EUT

"… there is no good scientific reason why it (prospect theory) should not replace expected utility in current research, and be given prominent space in economics textbooks."

– Colin F. Camerer
in "Prospect Theory in the wild: Evidence from the Field", 1998

Outline

- Cumulative Prospect Theory (CPT)
- **CPT Equilibrium Concepts - Nash and Correlated equilibrium**
- Results on the Geometry of CPT Equilibrium Notions
- Learning in CPT Games

Game Setup

$$\Gamma = (N, (A_i)_{i \in N}, (x_i)_{i \in N})$$

$N = \{1, 2, \ldots, n\}$ — Set of players

$a_i \in A_i$ — Actions for player i

$x_i : \prod_j A_j \to \mathbb{R}$ — Payoff function for player i

$a = (a_1, \ldots, a_n)$ — Action profile

$a \in A = \prod_i A_i$ — Set of Action profile

$a_{-i} \in A_{-i} = \prod_{j \neq i} A_j$ — Set of Action profile of opponents

Game Setup (EUT)

$N = \{1, 2, \ldots, n\}$ Set of players

$a_i \in A_i$ Actions for player i

$x_i : \prod_j A_j \to \mathbb{R}$ Payoff function for player i

For each player $i \in N$

Utility function

$u_i(\cdot)$

↓

Expected Utility

$U_i(L)$

Game Setup (CPT)

$N = \{1, 2, \ldots, n\}$ Set of players

$a_i \in A_i$ Actions for player i

$x_i : \prod_j A_j \to \mathbb{R}$ Payoff function for player i

For each player $i \in N$

Reference point
r_i

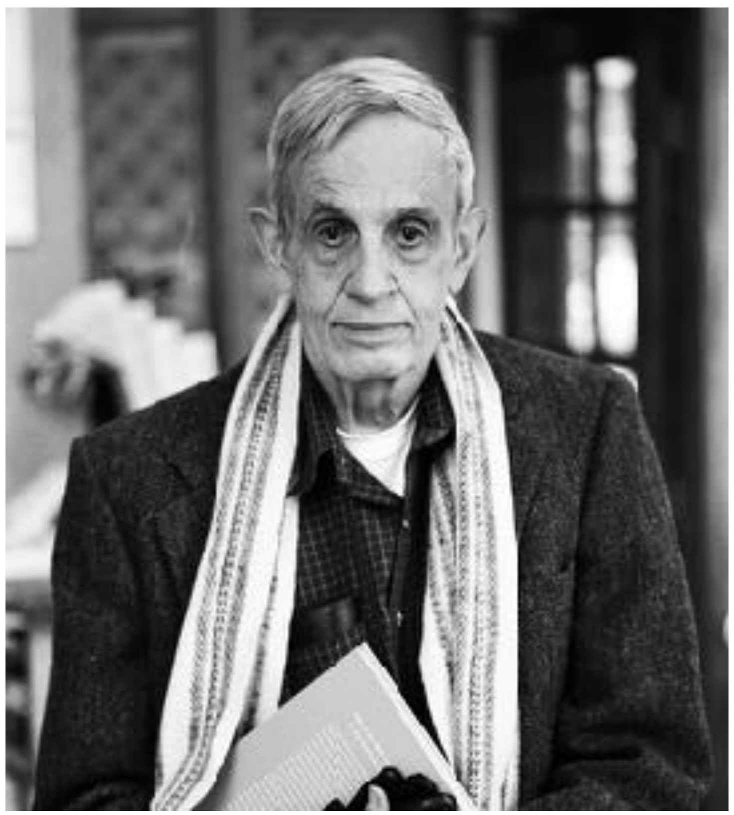
Value function
$v_i(\cdot)$

Probability weighting function
$w_i^{\pm}(\cdot)$

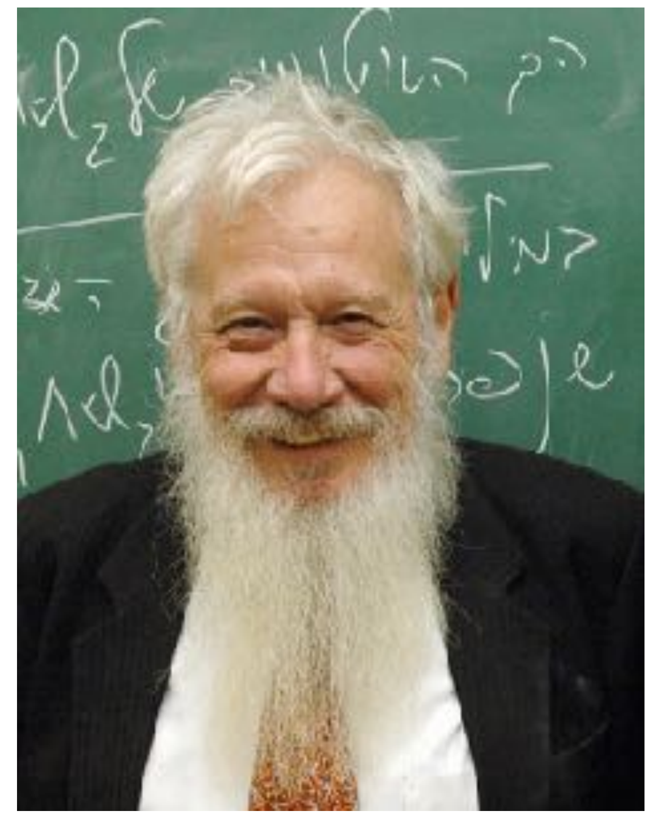
CPT value function

$V_i(L)$

Strategic Behavior in Games

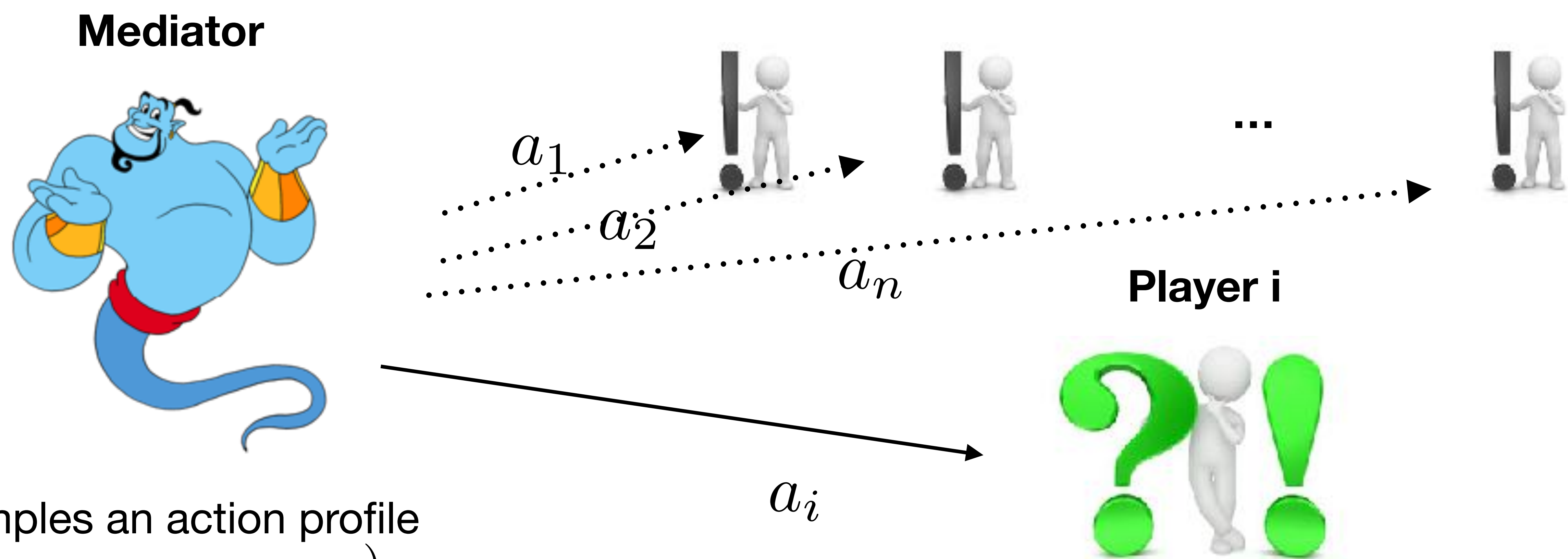


John Nash (1928–2015)
Nash Equilibrium



Robert Aumann (b. 1930)
Correlated Equilibrium

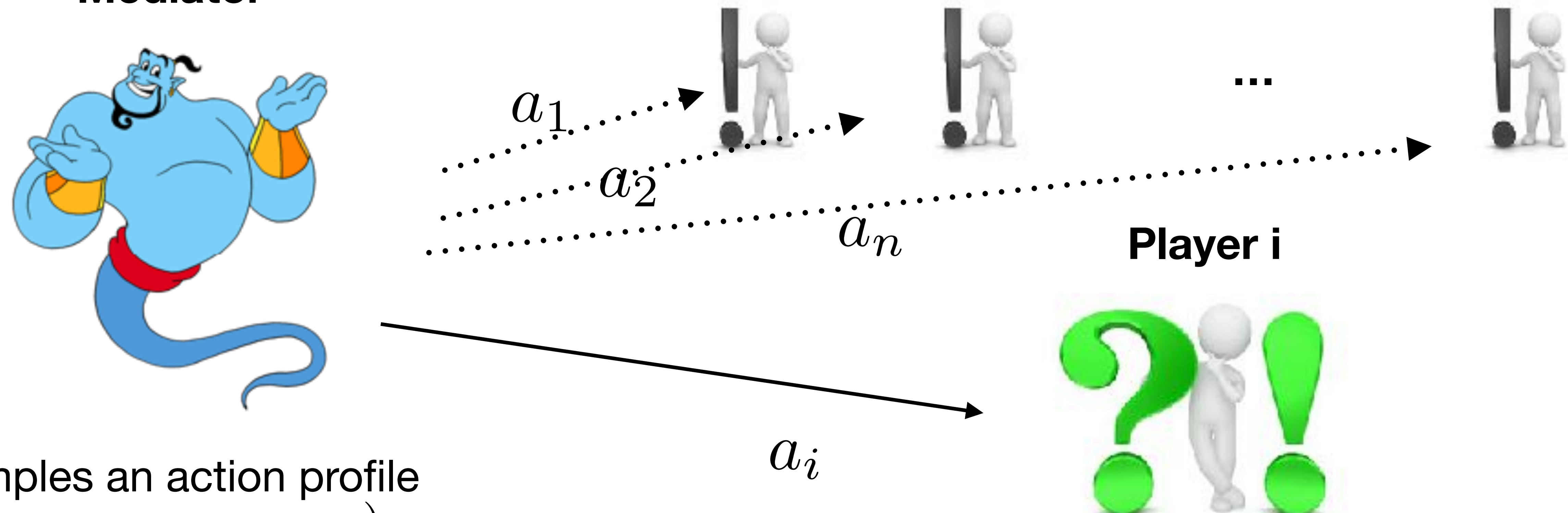
Correlated Equilibrium (CE)



Samples an action profile
$(a_1, \ldots, a_i, \ldots, a_n)$
from a distribution
$\mu \in \Delta(A)$

Correlated Equilibrium (CE)

Mediator



Samples an action profile
$(a_1, \ldots, a_i, \ldots, a_n)$
from a distribution
$\mu \in \Delta(A)$

Faces a lottery
corresponding to each action

$$L_i(\mu, a_i, \tilde{a}_i) := \left\{ \left(\underbrace{\nu(a_{-i})}, x_i(\tilde{a}_i, a_{-i}) \right) \right\}_{a_{-i} \in A_{-i}}$$

$$\nu(a_{-i}) = \frac{\mu(a_i, a_{-i})}{\sum_{a_{-i} \in A_{-i}} \mu(a_i, a_{-i})}$$

EUT Correlated Equilibrium (EUT CE)

Definition (Aumann 1987)
A distribution μ ϵ Δ(A) is an **(EUT) Correlated Equilibrium** if no player with EUT preferences has an incentive to deviate from his signaled action, i.e.

$$U_i(L_i(\mu, a_i, a_i)) \geq U_i(L_i(\mu, a_i, \tilde{a}_i))$$

for all $i, a_i, \tilde{a}_i$ such that $\mu_i(a_i) > 0$

EUT Correlated Equilibrium (EUT CE)

Definition (Aumann 1987)
A distribution μ ϵ Δ(A) is an **(EUT) Correlated Equilibrium** if no player with EUT preferences has an incentive to deviate from his signaled action, i.e.

$$U_i(L_i(\mu, a_i, a_i)) \geq U_i(L_i(\mu, a_i, \tilde{a}_i))$$

for all $i, a_i, \tilde{a}_i$ such that $\mu_i(a_i) > 0$

Incentive Constraints

EUT Correlated Equilibrium (EUT CE)

Definition (Aumann 1987)
A distribution μ ϵ Δ(A) is an **(EUT) Correlated Equilibrium** if no player with EUT preferences has an incentive to deviate from his signaled action, i.e.

$$U_i(L_i(\mu, a_i, a_i)) \geq U_i(L_i(\mu, a_i, \tilde{a}_i))$$

for all $i, a_i, \tilde{a}_i$ such that $\mu_i(a_i) > 0$

Incentive Constraints

Denote the set of all correlated equilibria by $C_{EUT}(\Gamma)$

CPT Correlated Equilibrium (CPT CE)

Definition (Keskin 2017)
A distribution μ ϵ Δ(A) is a **CPT Correlated Equilibrium** if no player with CPT preferences has an incentive to deviate from his signaled action, i.e.

$$V_i(L_i(\mu, a_i, a_i)) \geq V_i(L_i(\mu, a_i, \tilde{a}_i))$$

for all $i, a_i, \tilde{a}_i$ such that $\mu_i(a_i) > 0$

Incentive Constraints

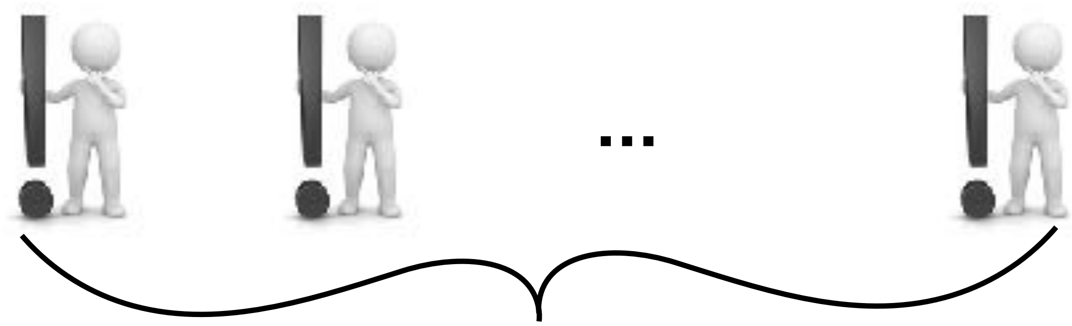
Denote the set of all CPT correlated equilibria by $C(\Gamma)$

Nash Equilibrium (NE)

$$\mu \in \Delta^*(A) = \{\mu \in \Delta(A) : \mu(a) = \mu_1(a_1) \times \cdots \times \mu_n(a_n) \quad \forall\, a \in A\}$$

Product form

μ_1 μ_2 μ_n



$$\mu_{-i} = \prod_{j \neq i} \mu_j$$

Player i

μ_i

$$L_i(\mu_{-i}, a_i) := \{(\mu_{-i}(a_{-i}), x_i(a_i, a_{-i}))\}_{a_{-i} \in A_{-i}}$$

EUT Nash Equilibrium (EUT NE)

Best response of player i to a product distribution μ ϵ Δ*(A)

$$BR_i(\mu) := \left\{ \mu_i^* \in \Delta(A_i) \,|\, \text{supp}(\mu_i^*) \subset \arg\max_{a_i \in A_i} U_i(L_i(\mu_{-i}, a_i)) \right\}$$

μ_i^* **Assigns positive probability only to optimal actions**

EUT Nash Equilibrium (EUT NE)

Best response of player i to a product distribution μ ϵ Δ*(A)

$$BR_i(\mu) := \left\{ \mu_i^* \in \Delta(A_i) | \operatorname{supp}(\mu_i^*) \subset \arg\max_{a_i \in A_i} U_i(L_i(\mu_{-i}, a_i)) \right\}$$

Definition (Nash 1951)

A product distribution μ ϵ Δ*(A) is EUT Nash equilibrium if

$$\mu^* \in BR_i(\mu^*) \text{ for all } i$$

EUT Nash Equilibrium (EUT NE)

Best response of player i to a product distribution μ ϵ Δ*(A)

$$BR_i(\mu) := \left\{ \mu_i^* \in \Delta(A_i) \middle| \operatorname{supp}(\mu_i^*) \subset \arg\max_{a_i \in A_i} U_i(L_i(\mu_{-i}, a_i)) \right\}$$

Definition (Nash 1951)

A product distribution μ ϵ Δ*(A) is EUT Nash equilibrium if

$$\mu^* \in BR_i(\mu^*) \text{ for all } i$$

Existence guaranteed by Kakutani fixed point theorem

CPT Nash Equilibrium (CPT NE)

Best response of player i to a product distribution μ ϵ Δ*(A)

$$BR_i(\mu) := \left\{ \mu_i^* \,\middle|\, \mathrm{supp}(\mu_i^*) \subset \arg\max_{a_i \in A_i} V_i\left(L_i(\mu_{-i}, a_i)\right) \right\}$$

<u>Definition (Keskin 2017)</u>

A product distribution μ ϵ Δ*(A) is CPT Nash equilibrium if

$$\mu^* \in BR_i(\mu^*) \text{ for all } i$$

CPT Nash Equilibrium (CPT NE)

Best response of player i to a product distribution μ ϵ Δ*(A)

$$BR_i(\mu) := \left\{ \mu_i^* \,\middle|\, \mathrm{supp}(\mu_i^*) \subset \arg\max_{a_i \in A_i} V_i\left(L_i(\mu_{-i}, a_i)\right) \right\}$$

<u>Definition (Keskin 2017)</u>

A product distribution μ ϵ Δ*(A) is CPT Nash equilibrium if

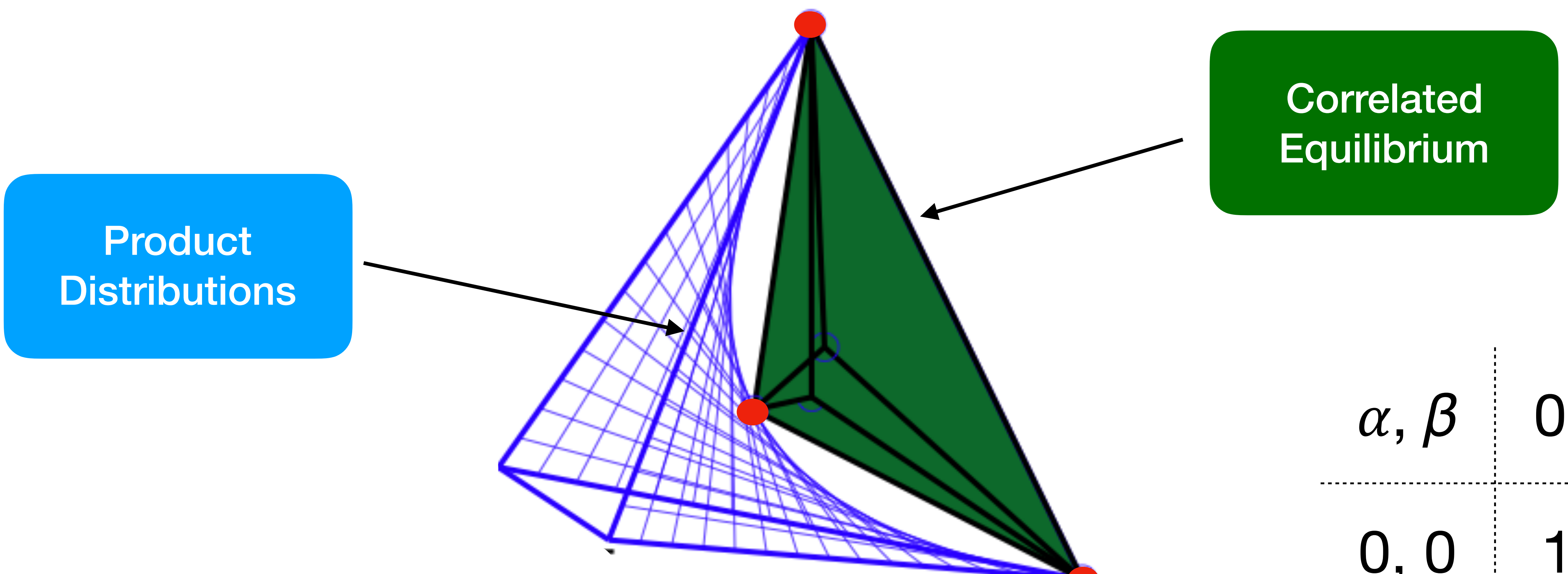
$$\mu^* \in BR_i(\mu^*) \text{ for all } i$$

Existence guaranteed by Kakutani fixed point theorem

Outline

- Cumulative Prospect Theory (CPT)
- CPT Equilibrium Concepts - Nash and Correlated equilibrium
- **Results on the Geometry of CPT Equilibrium Notions**
- Learning in CPT Games

Geometric Properties of Equilibria



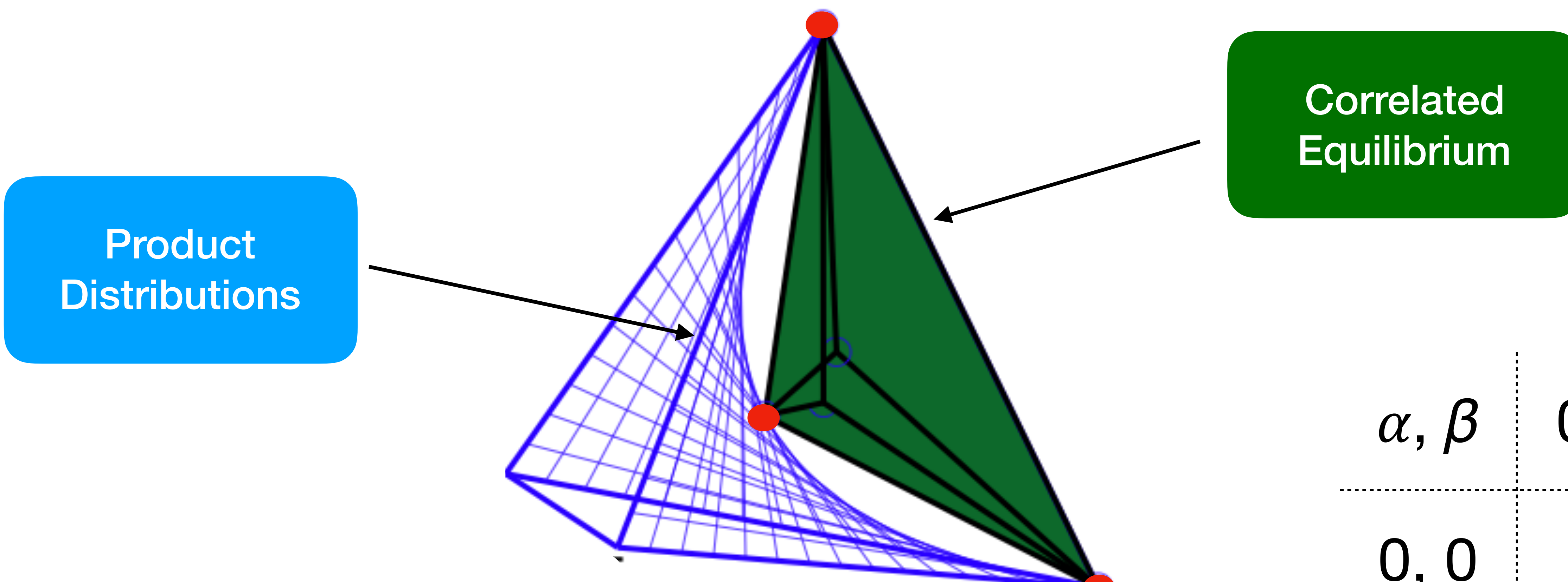
Representative picture for 2x2 games

α, β	0, 0
0, 0	1,1

Coordination 2x2 game

Geometric Properties of Equilibria

$$NE_{EUT}(\Gamma) = \Delta^*(A) \cap C_{EUT}(\Gamma)$$



Representative picture for 2x2 games

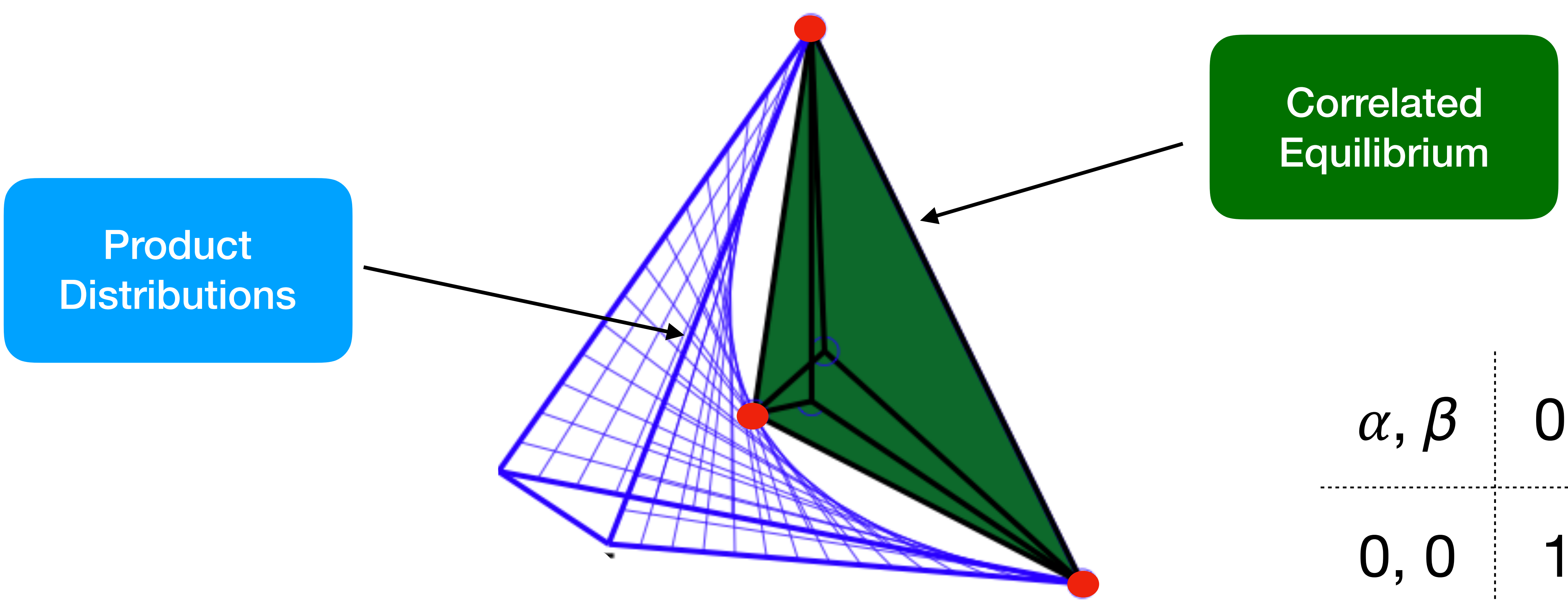
α, β	0, 0
0, 0	1,1

Coordination 2x2 game

Geometric Properties of Equilibria

$$NE_{EUT}(\Gamma) = \Delta^*(A) \cap C_{EUT}(\Gamma)$$

$$NE_{CPT}(\Gamma) = \Delta^*(A) \cap C(\Gamma)$$



Representative picture for 2x2 games

α, β	0, 0
0, 0	1,1

Coordination 2x2 game

Under EUT, Linearity in Probability helps!

Expanding Incentive Constraints using $U_i(L) = \Sigma\, u_i(x) p_i$

$$\sum_{a_{-i} \in A_{-i}} \mu(a) \left(u_i \left(x_i(a_i, a_{-i}) \right) - u_i \left(x_i(\tilde{a}_i, a_{-i}) \right) \right) \geq 0$$

$$\text{for all } i, a_i, \tilde{a}_i \in A_i$$

Under EUT, Linearity in Probability helps!

Expanding Incentive Constraints using $U_i(L) = \Sigma\, u_i(x) p_i$

$$\sum_{a_{-i} \in A_{-i}} \mu(a) \left(u_i \left(x_i(a_i, a_{-i}) \right) - u_i \left(x_i(\tilde{a}_i, a_{-i}) \right) \right) \geq 0$$

$$\text{for all } i, a_i, \tilde{a}_i \in A_i$$

The set $C_{EUT}(\Gamma)$ is a convex polytope

Under EUT, Linearity in Probability helps!

Expanding Incentive Constraints using $U_i(L) = \Sigma\, u_i(x) p_i$

$$\sum_{a_{-i} \in A_{-i}} \mu(a) \left(u_i \left(x_i(a_i, a_{-i})\right) - u_i \left(x_i(\tilde{a}_i, a_{-i})\right)\right) \geq 0$$

$$\text{for all } i, a_i, \tilde{a}_i \in A_i$$

The set $C_{EUT}(\Gamma)$ is a convex polytope

What happens under CPT?

CPT Example with Non-convex C(Γ)

Player 2

Player 1

	I	II	III	IV
0	2β,1	β+1,1	0,1	1,1
1	1.99,0	1.99,0	1.99,0	1.99,0

$$\beta = 1/w_1^+(0.5) = 2.299$$

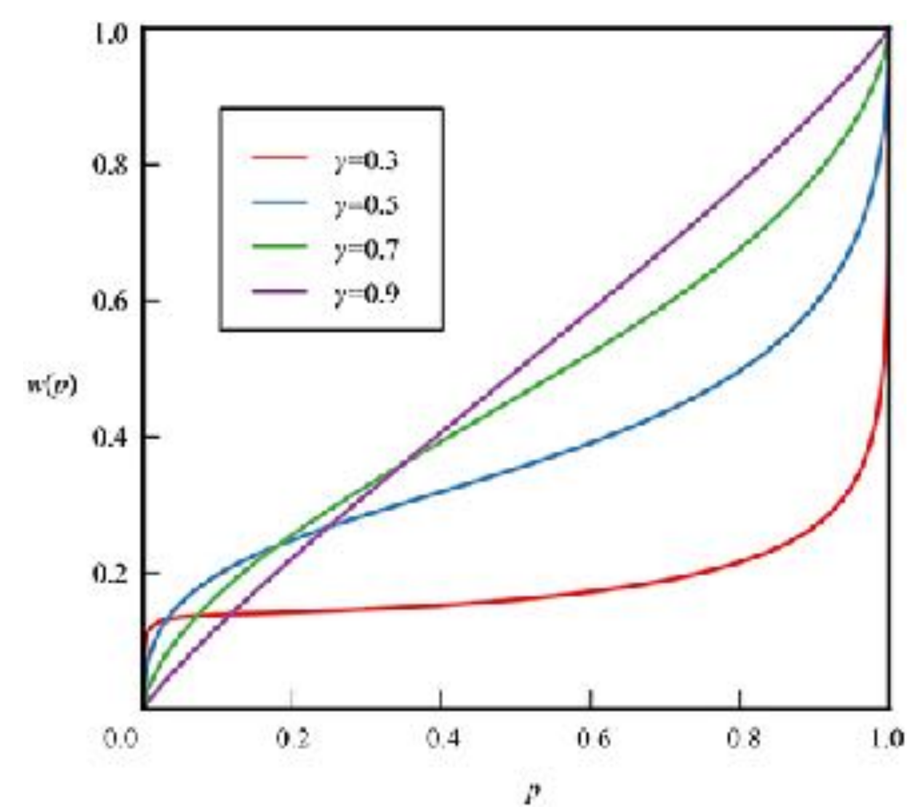
$$r_1 = r_2 = 0$$

$$\gamma_1 = 0.5, \gamma_2 = 1$$

$$v_1(z) = v_2(z) = z$$

Prelec 1998

$$w_i^+(p) = \exp\{-(-\ln p)^{\gamma_i}\}$$



CPT Example with Non-convex C(Γ)

Player 2

Player 1

	I	II	III	IV
0	2β,1	β+1,1	0,1	1,1
1	1.99,0	1.99,0	1.99,0	1.99,0

CPT Example with Non-convex C(Γ)

Player 2

Player 1

	I	II	III	IV
0	2β,1	β+1,1	0,1	1,1
1	1.99,0	1.99,0	1.99,0	1.99,0

Action 0 is player 1's Best Response to $\begin{cases} \mu_{odd} = (0.5, 0, 0.5, 0) \\ \mu_{even} = (0, 0.5, 0, 0.5) \end{cases}$

CPT Example with Non-convex C(Γ)

Player 2

Player 1

	I	II	III	IV
0	2β,1	β+1,1	0,1	1,1
1	1.99,0	1.99,0	1.99,0	1.99,0

Action 0 is player 1's Best Response to $\begin{cases} \mu_{odd} = (0.5, 0, 0.5, 0) \\ \mu_{even} = (0, 0.5, 0, 0.5) \end{cases}$

Action 1 is player 1's Best Response to $\mu_{unif} = (0.25, 0.25, 0.25, 0.25)$

CPT Example with Non-convex C(Γ)

Player 2

Player 1

	I	II	III	IV
0	2β,1	β+1,1	0,1	1,1
1	1.99,0	1.99,0	1.99,0	1.99,0

Action 0 is player 1's Best Response to $\begin{cases} \mu_{odd} = (0.5, 0, 0.5, 0) \\ \mu_{even} = (0, 0.5, 0, 0.5) \end{cases}$

Action 1 is player 1's Best Response to $\mu_{unif} = (0.25, 0.25, 0.25, 0.25)$

The set C(Γ) is Non-convex!

CPT Example

Player 2

Player 1

	I	II	III	IV
0	2β,1	β+1,1	0,1	1,1
1	1.99,0	1.99,0	1.99,0	1.99,0

$$\beta = 1/w_1^+(0.5) = 2.299$$

$$r_1 = r_2 = 0$$

$$\gamma_1 = 0.5, \gamma_2 = 1$$

$$v_1(z) = v_2(z) = z$$

Action 0 is player 1's Best Response to $\begin{cases} \mu_{odd} = (0.5, 0, 0.5, 0) \\ \mu_{even} = (0, 0.5, 0, 0.5) \end{cases}$

Action 1 is player 1's Best Response to $\mu_{unif} = (0.25, 0.25, 0.25, 0.25)$

$$V_1(L_1(\mu_{odd}, 0)) = 2\beta w_1^+(0.5) = 2, \qquad V_1(L_1(\mu_{odd}, 1)) = 1.99$$

$$V_1(L_1(\mu_{even}, 0)) = 1 + \beta w_1^+(0.5) = 2, \qquad V_1(L_1(\mu_{even}, 1)) = 1.99$$

Our results on the structure of CE (CPT)

<u>Result (P., Anantharam 2017)</u>
For any 2x2 game, the set C(Γ) is a **convex polytope**.

Our results on the structure of CE (CPT)

<u>Result (P., Anantharam 2017)</u>
For any 2x2 game, the set C(Γ) is a **convex polytope**.

<u>Result (P., Anantharam 2017)</u>
We provide an example of a 3x3 game for which the set C(Γ) is **disconnected**.

NE and CE (EUT)

$$NE_{EUT}(\Gamma) = \Delta^*(A) \cap C_{EUT}(\Gamma)$$

Theorem (Nau et al 2003)
The Nash equilibria all lie on the boundary of the correlated equilibria polytope.

NE and CE (EUT)

$$NE_{EUT}(\Gamma) = \Delta^*(A) \cap C_{EUT}(\Gamma)$$

Theorem (Nau et al 2003)
The Nash equilibria all lie on the boundary of the correlated equilibria polytope.

Comments:

NE and CE (EUT)

$$NE_{EUT}(\Gamma) = \Delta^*(A) \cap C_{EUT}(\Gamma)$$

Theorem (Nau et al 2003)
The Nash equilibria all lie on the boundary of the correlated equilibria polytope.

Comments:

1. Boundary of the correlated equilibrium set when it is viewed as a subset of Δ(A).

NE and CE (EUT)

$$NE_{EUT}(\Gamma) = \Delta^*(A) \cap C_{EUT}(\Gamma)$$

Theorem (Nau et al 2003)
The Nash equilibria all lie on the boundary of the correlated equilibria polytope.

Comments:

1. Boundary of the correlated equilibrium set when it is viewed as a subset of Δ(A).
2. If the set of correlated equilibria is not "full dimensional" then the statement is trivial.

NE and CE (EUT)

$$NE_{EUT}(\Gamma) = \Delta^*(A) \cap C_{EUT}(\Gamma)$$

Theorem (Nau et al 2003)
The Nash equilibria all lie on the boundary of the correlated equilibria polytope.

Comments:

1. Boundary of the correlated equilibrium set when it is viewed as a subset of Δ(A).
2. If the set of correlated equilibria is not “full dimensional” then the statement is trivial.
3. The statement cannot be strengthened because in the case of less than full dimensional C(Γ), Nash equilibria can lie in the relative interior of this convex polytope.

NE and CE (EUT)

<u>Theorem (Nau et al 2003)</u>
The Nash equilibria all lie on the boundary of the correlated equilibria polytope.

Proof sketch:
A Nash equilibria renders every player indifferent among all of her own strategies, hence it satisfies all of the incentive constraints with **equality**, at least one of which is non trivial if the game is non trivial, and hence lies on one of the faces of the convex polytope CE_{EUT}.

$$\sum_{s_{-i} \in S_{-i}} \mu(s)(u_i(h_i(s_i, s_{-i})) - u_i(h_i(d_i, s_{-i}))) \geq 0,$$

$$\text{for all } i \text{ and for all } s_i, d_i \in S_i$$

NE and CE (CPT)

$$NE_{CPT}(\Gamma) = \Delta^*(A) \cap C(\Gamma)$$

Theorem (P., Anantharam 2017)
The CPT Nash equilibria all lie on the boundary of the CPT correlated equilibria set.

NE and CE (CPT)

$$NE_{CPT}(\Gamma) = \Delta^*(A) \cap C(\Gamma)$$

Theorem (P., Anantharam 2017)
The CPT Nash equilibria all lie on the boundary of the CPT correlated equilibria set.

Proof sketch

NE and CE (CPT)

$$NE_{CPT}(\Gamma) = \Delta^*(A) \cap C(\Gamma)$$

Theorem (P., Anantharam 2017)
The CPT Nash equilibria all lie on the boundary of the CPT correlated equilibria set.

Proof sketch
1. Only need to focus on completely mixed CPT NE.

NE and CE (CPT)

$$NE_{CPT}(\Gamma) = \Delta^*(A) \cap C(\Gamma)$$

Theorem (P., Anantharam 2017)
The CPT Nash equilibria all lie on the boundary of the CPT correlated equilibria set.

Proof sketch
1. Only need to focus on completely mixed CPT NE.
2. CPT Nash equilibrium satisfies all of the incentive constraints with equality.

NE and CE (CPT)

$$NE_{CPT}(\Gamma) = \Delta^*(A) \cap C(\Gamma)$$

Theorem (P., Anantharam 2017)
The CPT Nash equilibria all lie on the boundary of the CPT correlated equilibria set.

Proof sketch
1. Only need to focus on completely mixed CPT NE.
2. CPT Nash equilibrium satisfies all of the incentive constraints with equality.
3. There exists a direction in the probability space along which at least one incentive constraint is violated.

NE and CE (CPT)

$$NE_{CPT}(\Gamma) = \Delta^*(A) \cap C(\Gamma)$$

Theorem (P., Anantharam 2017)
The CPT Nash equilibria all lie on the boundary of the CPT correlated equilibria set.

Proof sketch

1. Only need to focus on completely mixed CPT NE.
2. CPT Nash equilibrium satisfies all of the incentive constraints with equality.
3. There exists a direction in the probability space along which at least one incentive constraint is violated.
4. Thus CPT Nash equilibria cannot have a ball around it that is completely contained inside the set of CPT correlated equilibria.

A Useful Lemma

Two distinct lotteries (let $p_j > 0$ for j=1,…,t)

$L_1 =$

probability	p_1	p_2	.	.	.	p_t
outcome	x_1	x_2	.	.	.	x_t

$L_2 =$

probability	p_1	p_2	.	.	.	p_t
outcome	y_1	y_2	.	.	.	y_t

that satisfy **either** of the two properties

1. they are not similarly ranked or
2. neither of them dominates the other

then ∃ a direction $\delta = (\delta_1, \ldots, \delta_n)$, $\Sigma\, \delta_i = 0$ such that

$$V(p + \epsilon\delta, x) - V(p + \epsilon\delta, y) < V(p, x) - V(p, y)$$

Outline

- Cumulative Prospect Theory (CPT)
- CPT Equilibrium Concepts - Nash and Correlated equilibrium
- Results on the Geometry of CPT Equilibrium Notions
- **Learning in CPT Games**

Learning in Games

Learning in Games

- Neoclassical economics:
  - hyper-rational players,
  - completely understand the structure of the game,
  - have a coherent model of others' behavior,
  - make rational calculations of infinite complexity,
  - and all of this is common knowledge

Learning in Games

- Neoclassical economics:
    - hyper-rational players,
    - completely understand the structure of the game,
    - have a coherent model of others' behavior,
    - make rational calculations of infinite complexity,
    - and all of this is common knowledge

- Learning in Games:
    - boundedly-rational players,
    - players make decisions on limited data,
    - and use simple predictive models

Learning in Games

Learning in Games

- Classical solution concepts (like NE and CE) can be “recovered” via this route.

Learning in Games

- Classical solution concepts (like NE and CE) can be "recovered" via this route.

- Examples:
  - Fictitious play: Nash Equilibrium in zero sum games, potential games, 2x2 games
  - **Calibrated learning**, Adaptive heuristics, Follow the perturbed leader: Correlated Equilibrium

Learning in Games

- Classical solution concepts (like NE and CE) can be "recovered" via this route.

- Examples:
    - Fictitious play: Nash Equilibrium in zero sum games, potential games, 2x2 games
    - **Calibrated learning**, Adaptive heuristics, Follow the perturbed leader: Correlated Equilibrium

- **Question: What if the players behave according to CPT?**

Repeated games

- A Game Γ is played repeatedly at each step t = 1,2,…
- Player i's action sequence: $a_i^1, a_i^2, \ldots$
- Action profile at step t: $a^t = (a_1^t, \ldots, a_n^t)$
- History at step t: $H^t = (a^1, a^2, \ldots, a^{t-1})$
- Randomized strategy sequence for player i at step t:

$$\sigma_i^t : H^t \to \Delta(A_i)$$

- Empirical distribution at step t: ξ^t

$$\xi^t(a) = \frac{\#\text{action profile } a \text{ appears in } H^t}{t-1}$$

Foster Vohra result

Foster

Vohra

- At every step t, each player i predicts a distribution $\mu^t_{-i} \in \Delta(A_{-i})$ on the action profile of the other players.

- Based on this prediction she plays a best EUT response a^t_i

Foster Vohra result

Foster

Vohra

- At every step t, each player i predicts a distribution $\mu_{-i}^t \in \Delta(A_{-i})$ on the action profile of the other players.
- Based on this prediction she plays a best EUT response a_i^t

Theorem (Foster, Vohra 1997)

If each players' predictions are calibrated, then the Empirical Distribution of play convergence to the set of **correlated equilibria**.

$$\lim_{t\to\infty} d(\xi^t, C_{EUT}(\Gamma)) = 0$$

Calibrated Prediction

Nature vs Forecaster

Step	1	2	3	4	5	6	7	8	...
Nature	0	1	1	0	0	0	1	0	...
Forecaster	10%	70%	80%	30%	10%	10%	80%	30%	...

Calibrated Prediction

Nature vs Forecaster

Step	1	2	3	4	5	6	7	8	...
Nature	0	1	1	0	0	0	1	0	...
Forecaster	10%	70%	80%	30%	10%	10%	80%	30%	...

Nature: $y^1, y^2, \ldots \in S$.

Forecaster: $q^1, q^2, \ldots \in \Delta(S)$.

Calibrated Prediction

Nature vs Forecaster

Step	1	2	3	4	5	6	7	8	...
Nature	0	1	1	0	0	0	1	0	...
Forecaster	10%	70%	80%	30%	10%	10%	80%	30%	...

Nature: $y^1, y^2, \ldots \in S$.

Forecaster: $q^1, q^2, \ldots \in \Delta(S)$.

$$N(q,t) = \#\ \text{Forecaster predicts q up to step } t$$

Calibrated Prediction

Nature vs Forecaster

Step	1	2	3	4	5	6	7	8	...
Nature	0	1	1	0	0	0	1	0	...
Forecaster	10%	70%	80%	30%	10%	10%	80%	30%	...

Nature: $y^1, y^2, \ldots \in S$.

Forecaster: $q^1, q^2, \ldots \in \Delta(S)$.

$$N(q,t) = \#\ \text{Forecaster predicts q up to step } t$$

$$\rho(q,y,t) = \frac{\#\text{Forecaster predicts } q \text{ and Nature plays } y \text{ up to step } t}{N(q,t)}$$

Calibrated Prediction

Nature vs Forecaster

Step	1	2	3	4	5	6	7	8	...
Nature	0	1	1	0	0	0	1	0	...
Forecaster	10%	70%	80%	30%	10%	10%	80%	30%	...

Nature: $y^1, y^2, \ldots \in S$.

Forecaster: $q^1, q^2, \ldots \in \Delta(S)$.

$$N(q,t) = \#\text{ Forecaster predicts q up to step } t$$

$$\rho(q,y,t) = \frac{\#\text{Forecaster predicts } q \text{ and Nature plays } y \text{ up to step } t}{N(q,t)}$$

$$\lim_{t\to\infty} \sum_{q\in Q^t} |\rho(q,y,t) - q(y)| \frac{N(q,t)}{t} = 0 \text{ for all } y \in S,$$

Example of CPT calibrated learning

Player 2/Player 3

Player 1

	I	II	III	IV
0	2β,1	β+1,1	0,1	1,1
1	1.99,0	1.99,0	1.99,0	1.99,0

- If Player 2 Action ≠ Player 3 Action then all players receive a payoff of -1.
- If Player 2 action = player 3 action then Player 1 receives first payoff shown in table and players 2 and 3 each receive the second payoff.

Non convergence of Calibrated learning to CPT correlated equilibrium

Action 0 is player 1's Best Response to $\begin{cases} \mu_{odd} = (0.5, 0, 0.5, 0) \\ \mu_{even} = (0, 0.5, 0, 0.5) \end{cases}$

Action 1 is player 1's Best Response to $\mu_{unif} = (0.25, 0.25, 0.25, 0.25)$

Step	1	2	3	4	5	6
Player 1 Action	0	0	0	0	0	0
Player 2 Action	I	II	III	IV	I	II
Player 3 Action	I	II	III	IV	I	II
Player 1 Forecast	μ_{odd}	μ_{even}	μ_{odd}	μ_{even}	μ_{odd}	μ_{even}

	I	II	III	IV
0	0.25	0.25	0.25	0.25
1	0	0	0	0

Structure of CE: Non-convexity is the issue

$C(\Gamma, i, a_i) \subset \Delta(A_{-i})$: distributions for which action a_i is player i's Best Response

Structure of CE: Non-convexity is the issue

$C(\Gamma, i, a_i) \subset \Delta(A_{-i})$: distributions for which action a_i is player i's Best Response

$$C(\Gamma, i) := \{\mu \in \Delta(A) | \mu(\cdot | a_i) \in C(\Gamma, i, a_i), \forall a_i \in \operatorname{supp}(\mu_i)\}$$

Distributions for which player i has no incentive to deviate

Structure of CE: Non-convexity is the issue

$C(\Gamma, i, a_i) \subset \Delta(A_{-i})$: distributions for which action a_i is player i's Best Response

$$C(\Gamma, i) := \{\mu \in \Delta(A) | \mu(\cdot | a_i) \in C(\Gamma, i, a_i), \forall a_i \in \operatorname{supp}(\mu_i)\}$$

Distributions for which player i has no incentive to deviate

$$C(\Gamma) = \cap_{i \in N} C(\Gamma, i)$$

Structure of CE: Non-convexity is the issue

$C(\Gamma, i, a_i) \subset \Delta(A_{-i})$: distributions for which action a_i is player i's Best Response

$$C(\Gamma, i) := \{\mu \in \Delta(A) | \mu(\cdot | a_i) \in C(\Gamma, i, a_i), \forall a_i \in \operatorname{supp}(\mu_i)\}$$

Distributions for which player i has no incentive to deviate

$$C(\Gamma) = \cap_{i \in N} C(\Gamma, i)$$

$$D(\Gamma) = \cap_{i \in N} co(C(\Gamma, i))$$

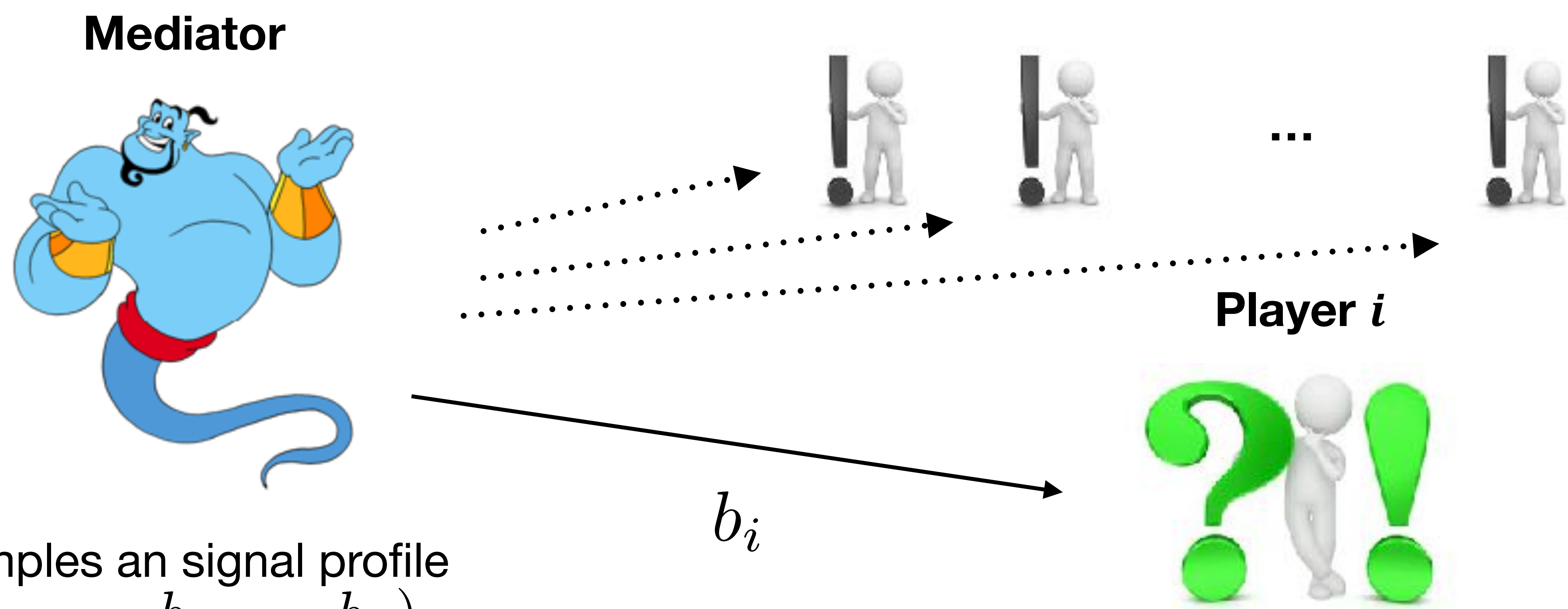
Mediated Games

$N = \{1, 2, \ldots, n\}$	Set of players
$a_i \in A_i$	Actions for player i
$x_i : \prod_j A_j \to \mathbb{R}$	Payoff function for player i

Mediated Games

$N = \{1, 2, \ldots, n\}$	Set of players
$a_i \in A_i$	Actions for player i
$x_i : \prod_j A_j \to \mathbb{R}$	Payoff function for player i
$b_i \in B_i$	Signal set for player i
$b = (b_1, \ldots, b_n)$	Signal profile
$B = \prod_i B_i$	Set of Signal profiles
$\sigma_i : B_i \to \Delta(A_i)$	Strategy of player i

Mediated Games



Samples an signal profile
$(b_1, \ldots, b_i, \ldots, b_n)$
from a distribution

$$\psi \in \Delta(B)$$

Mediated Games

Mediator

Samples an signal profile

$(b_1, \ldots, b_i, \ldots, b_n)$

from a distribution

$\psi \in \Delta(B)$

b_i

Player *i*

Faces a lottery
corresponding to each action

$$L_i(\psi, b_i, a_i) := \Big\{ \underbrace{\tilde{\mu}_{-i}(a_{-i}|b_i)}, x_i(a_i, a_{-i}) \Big\}_{a_{-i} \in A_{-i}}$$

$$\tilde{\mu}_{-i}(a_{-i}|b_i) := \sum_{b_{-i} \in B_{-i}} \psi_{-i}(b_{-i}|b_i) \prod_{j \in N \setminus i} \sigma_j(b_j)(a_j)$$

Mediated CPT NE

Best response of player i

$$BR_i(\psi,\sigma) := \left\{ \sigma_i^* : B^i \to \Delta(A_i) \middle| \text{ for all } b_i \in B_i \right.$$

$$\left. \operatorname{supp}(\sigma_i^*(b_i)) \subset \arg\max_{a_i \in A_i} V_i\left(L_i(\psi, b_i, a_i)\right) \right\}$$

$\sigma_i^*(b_i)$ **Assigns positive probability only to optimal actions**

Mediated CPT NE

Best response of player i

$$BR_i(\psi,\sigma) := \left\{ \sigma_i^* : B^i \to \Delta(A_i) \middle| \text{ for all } b_i \in B_i \right.$$

$$\left. \text{supp}(\sigma_i^*(b_i)) \subset \arg\max_{a_i \in A_i} V_i\left(L_i(\psi, b_i, a_i)\right) \right\}$$

Definition (P., Anantharam 2018)

A randomized strategy profile σ is **CPT Mediated Nash equilibrium** if

$$\sigma_i \in BR_i(\psi,\sigma) \text{ for all } i$$

Mediated CPT CE

Distribution induced by ψ and strategy profile σ on the set of action profiles

$$\eta(\psi, \sigma)(a) := \sum_{b \in B} \psi(b) \prod_{i \in N} \sigma_i(b_i)(a_i)$$

Mediated CPT CE

Distribution induced by ψ and strategy profile σ on the set of action profiles

$$\eta(\psi, \sigma)(a) := \sum_{b \in B} \psi(b) \prod_{i \in N} \sigma_i(b_i)(a_i)$$

Definition (P., Anantharam 2018)

A distribution μ ϵ Δ(A) is a **mediated CPT correlated equilibrium** iff there exists a signal system B_i, a mediator distribution ψ and a mediated CPT Nash equilibrium σ with respect to them such that

$$\eta(\psi, \sigma) = \mu$$

Mediated CPT CE

Theorem (P., Anantharam 2018)
The set of all Mediated CPT correlated equilibria is

$$D(\Gamma) = \cap_{i \in N} co(C(\Gamma, i))$$

Corollary 1
For EUT games, $D(\Gamma) = C(\Gamma)$.

Corollary 2
For 2x2 CPT games, $D(\Gamma) = C(\Gamma)$.

Convergence of Calibrated learning

- At every step t, each player i predicts a distribution $\mu_{-i}^t \in \Delta(A_{-i})$ on the action profile of the other players.

- Based on this prediction she plays a best CPT response a_i^t

<u>Theorem (P., Anantharam 2018)</u>
If each players' predictions are calibrated, then the Empirical Distribution of play convergence to the set of **mediated CPT correlated equilibria**.

$$\lim_{t \to \infty} d(\xi^t, D(\Gamma)) = 0$$

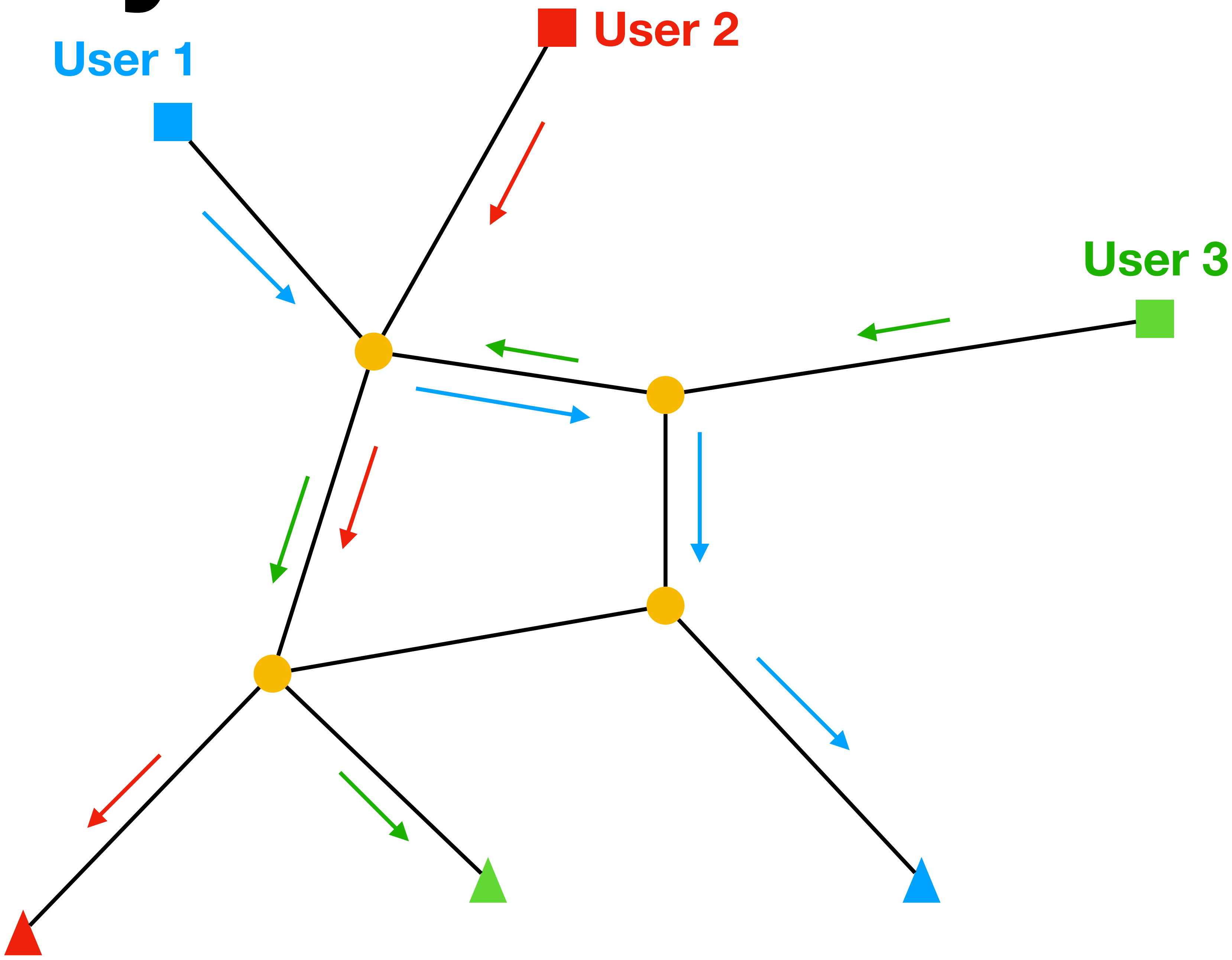
Converse

Theorem

If the sets $C(\Gamma, i, a_i)$ do not have any isolated points, then for any $\mu \in D(\Gamma)$ there exists a sequence of play and corresponding assessments that are calibrated such that the Empirical Distribution converges to μ.

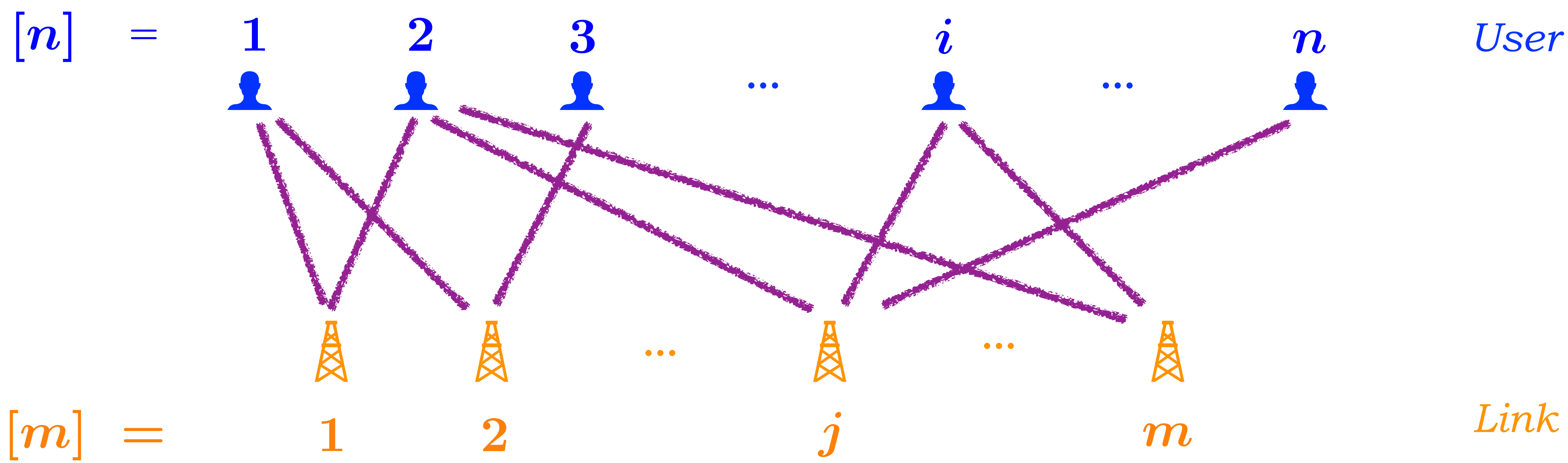
Such games are Generic

Kelly Network Setup

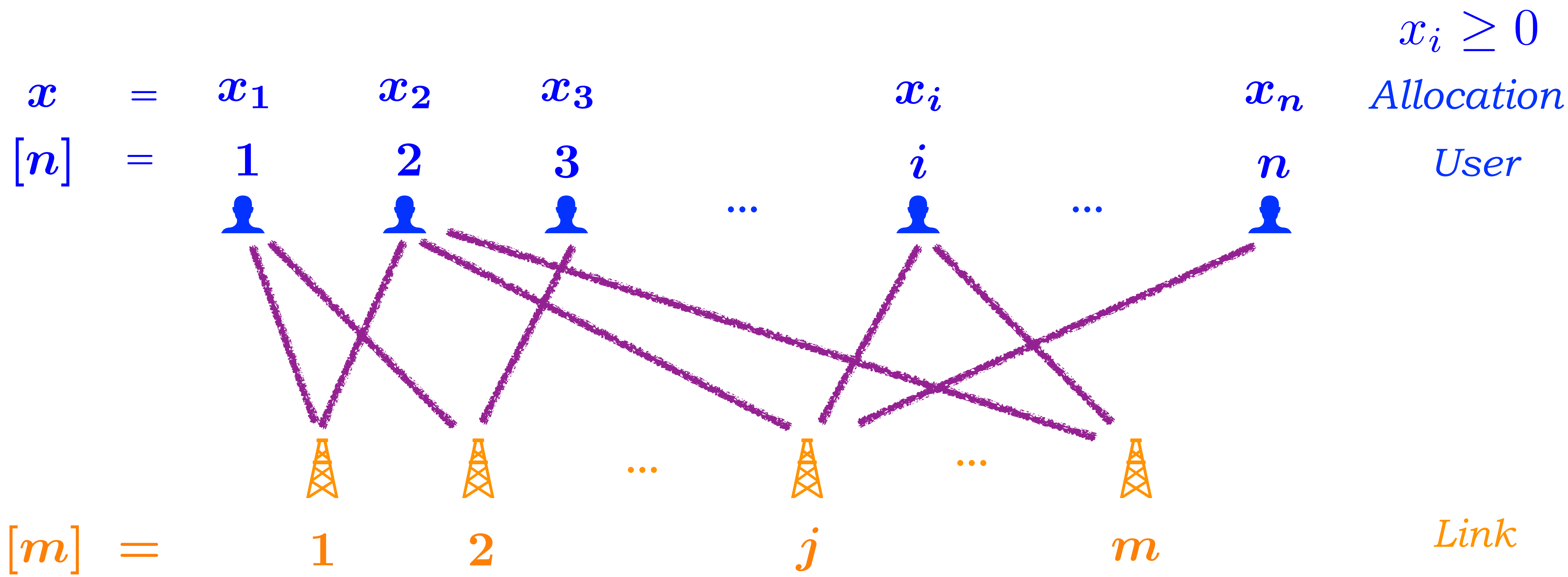


Kelly (1997) - Charging and rate control for elastic traffic

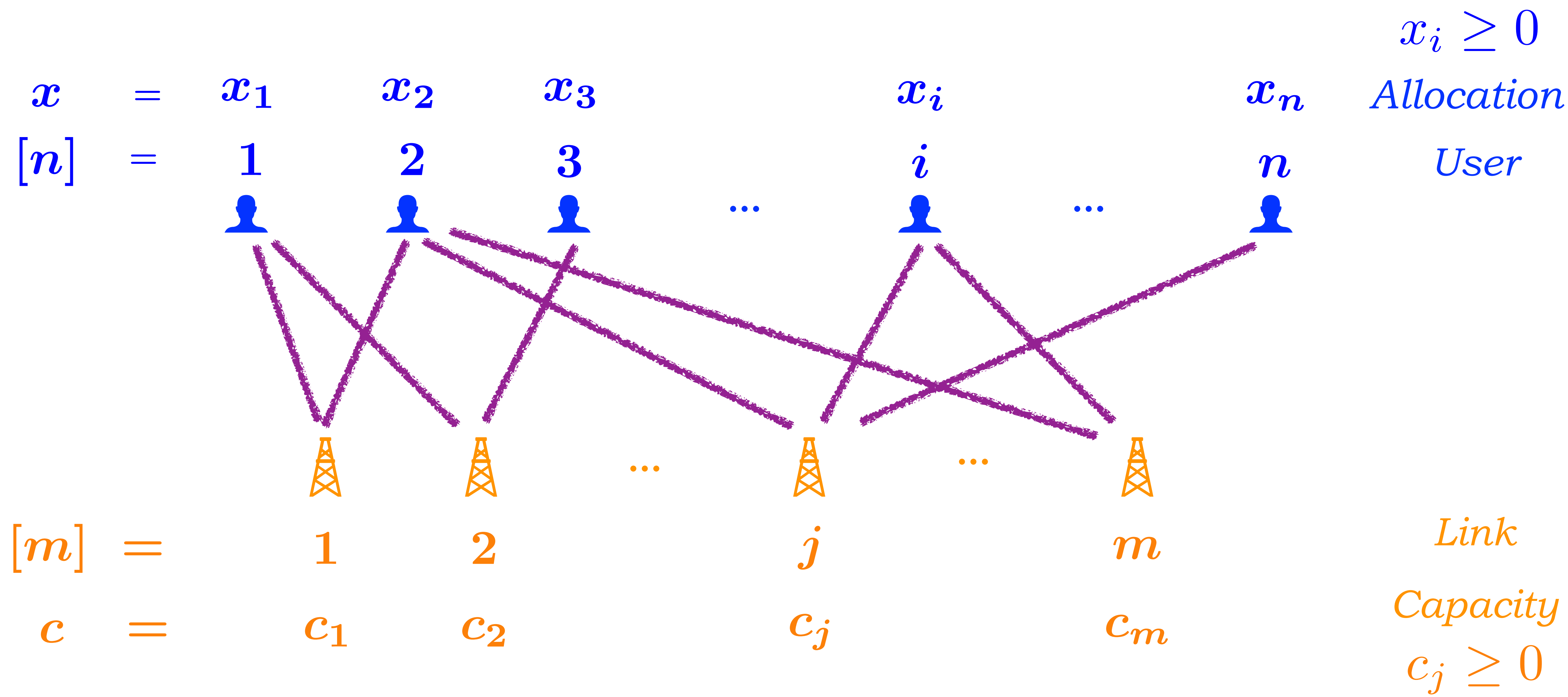
Kelly Network setup



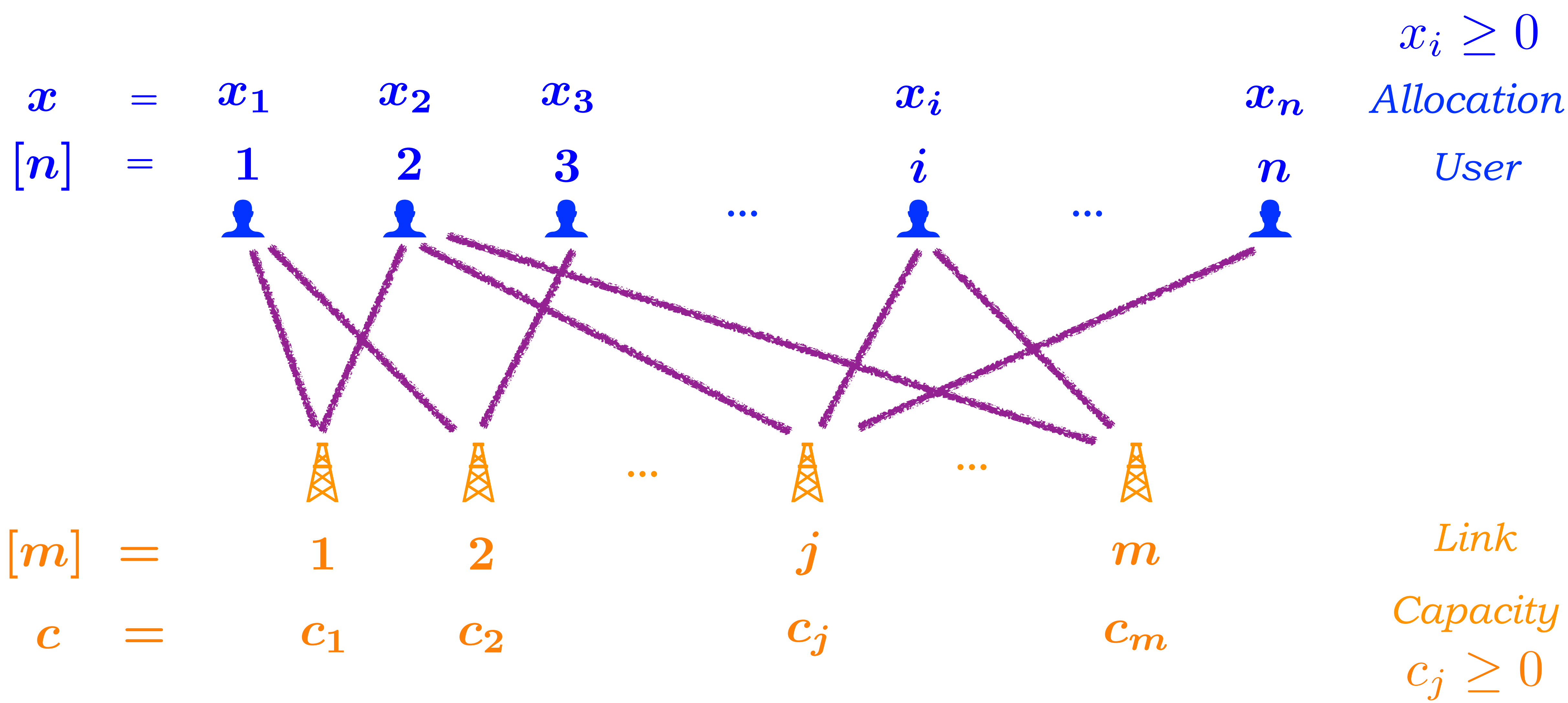
Kelly Network setup



Kelly Network setup



Kelly Network setup



$$\sum_{i \in R_j} x_i \leq c_j, \forall j$$

Link Constraints

System problem in the Kelly model

SYSTEM(U, A, C)

Maximize

$$\sum_{i=1}^{n} U_i(x_i)$$

subject to

$$\sum_{i \in R_j} x_i \leq c_j, \quad \forall j,$$

$$x_i \geq 0, \quad \forall i.$$

User problem in the Kelly model

$$\text{USER}_i(U_i, \lambda_i)$$

Maximize

$$U_i\left(\frac{w_i}{\lambda_i}\right) - w_i$$

subject to

$$w_i \geq 0.$$

w_i: Amount per unit time that user i is willing to pay

λ_i: charge per unit flow that the network presents to user i

Network problem in the Kelly model

$$\text{NETWORK}(A, C; w)$$

Maximize

$$\sum_i w_i \log x_i$$

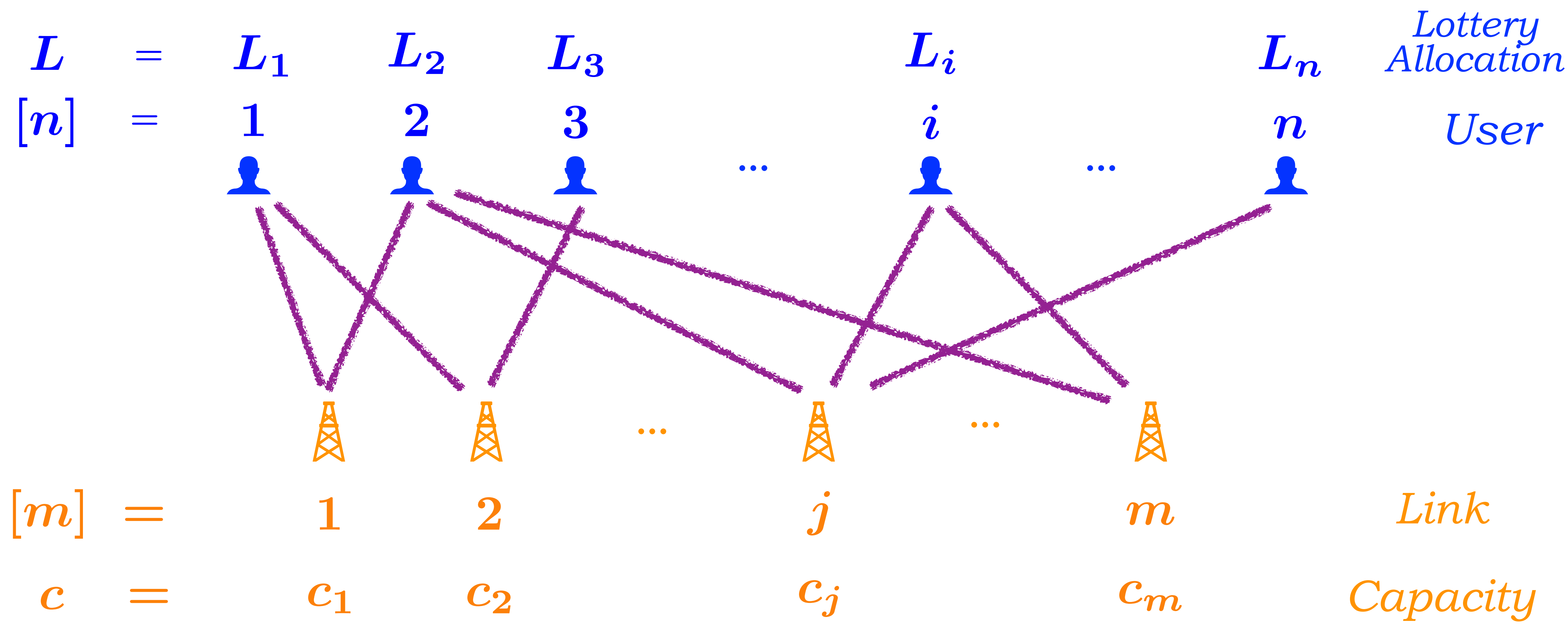
subject to

$$\sum_{i \in R_j} x_i \le c_j, \quad \forall j,$$

$$x_i \ge 0, \quad \forall i.$$

Since $w_i = \lambda_i x_i$ having found x_i the network can present λ_i to user i

Lottery Allocation



$L_i =$

probability	0.1	0.3	0.25	0.35
allocation	10	5	3	6

Implementable Allocation Schemes

User i

Alternative l

$y_i(l)$	1	2	3
1	10	3	2
2	5	5	5
3	3	10	2
4	0	5	10

$p(l)$
0.1
0.4
0.2
0.3

$$\sum_{i \in R_j} y_i(l) \leq c_j, \forall j, l$$

Implementable Allocation Schemes

User i

Alternative l

$y_i(l)$	1	2	3
1	10	3	2
2	5	5	5
3	3	10	2
4	0	5	10

$p(l)$
0.25
0.25
0.25
0.25

Uniformly Distributed

$$\sum_{i \in R_j} y_i(l) \leq c_j, \forall j, l$$

permutations and decision weights

User i (columns), Alternative l (rows)

$y_i(l)$	1	2	3
1	10	3	2
2	5	5	5
3	3	10	2
4	0	5	10

User i (columns), Alternative l (rows)

$z_i(l)$	1	2	3
1	10	10	10
2	5	5	5
3	3	5	2
4	0	3	2

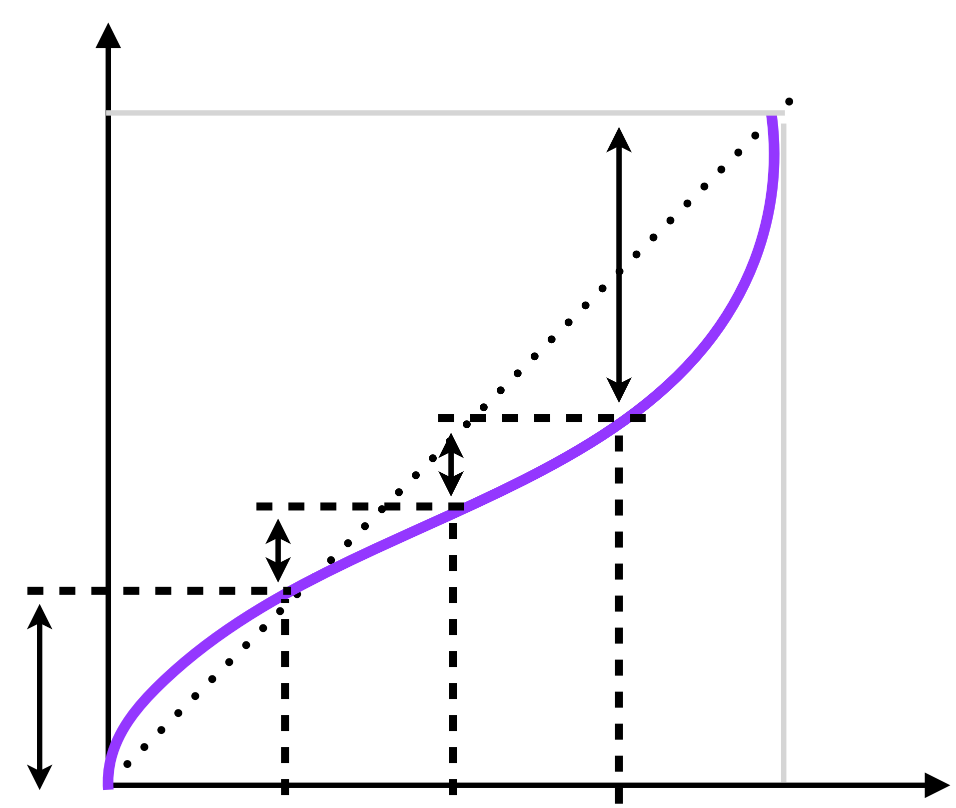
User i (columns), Alternative l (rows)

$\pi_i(l)$	1	2	3
1	1	4	3
2	2	2	2
3	3	1	4
4	4	3	1

$$y_i(l) = z_i(\pi_i(l))$$

$$V_i(L_i) = \sum_{l\in[k]} h_i(l) v_i(z_i(l))$$

$$w_i(\frac{l}{k}) - w_i(\frac{l-1}{k})$$



System Problem

$$\mathrm{SYS}[z, \pi; h, v, A, c]$$

Maximize $$\sum_{i}^{n} V_i(L_i)$$

Subject to $$\sum_{i \in R_j} y_i(l) \leq c_j, \forall j, l$$

System Problem

$$\mathrm{SYS}[z, \pi; h, v, A, c]$$

Maximize $$\sum_{i=1}^{n} \sum_{l=1}^{k} h_i(l) v_i(z_i(l))$$

Subject to $$\sum_{i \in R_j} y_i(l) \leq c_j, \forall j, l$$

System Problem

$$\mathrm{SYS}[z, \pi; h, v, A, c]$$

Maximize

$$\sum_{i=1}^{n} \sum_{l=1}^{k} h_i(l) v_i(z_i(l))$$

Subject to

$$\sum_{i \in R_j} z_i(\pi_i(l)) \leq c_j, \forall j, l$$

$$z_i(l) \geq z_i(l+1), \forall i, l$$

$$\pi_i \in S_k, \forall i$$

Fixed permutation problem

Maximize

$$\sum_{i=1}^{n} \sum_{l=1}^{k} h_i(l) v_i(z_i(l))$$

Subject to

$$\sum_{i \in R_j} z_i(\pi_i(l)) \leq c_j, \forall j, l$$

$$z_i(l) \geq z_i(l+1), \forall i, l$$

$$\pi_i \in S_k, \forall i$$

Fixed permutation problem

Maximize

$$\sum_{i=1}^{n} \sum_{l=1}^{k} h_i(l) v_i(z_i(l))$$

Subject to

$$\sum_{i \in R_j} z_i(\pi_i(l)) \leq c_j, \forall j, l$$

$$z_i(l) \geq z_i(l+1), \forall i, l$$

Fix

$$\pi_i \in S_k, \forall i$$

Fixed permutation problem

$$\mathrm{SYS\_FIX}[z;\pi,h,v,A,c]$$

Maximize $$\sum_{i=1}^{n}\sum_{l=1}^{k} h_i(l)v_i(z_i(l))$$

Subject to $$\sum_{i\in R_j} z_i(\pi_i(l)) \le c_j, \forall j, l$$

$$z_i(l) \ge z_i(l+1), \forall i, l$$

Convex Optimization

User pricing through menu of rates

l	Menu of Rates $r_i(l)$	Vector of Budgets $m_i(l)$	Incremental Allocation $\delta_i(l)$	Lottery Allocation $z_i(l)$
1	1	5	5 = 5/1	10 = 5 + 5
2	2	0	0 = 0/2	5 = 5 + 0
3	3	6	2 = 6/3	5 = 3 + 2
4	5	15	3 = 15/5	3

User problem

$$\mathrm{USER}[m_i; r_i, h_i, v_i]$$

Maximize $$\sum_{l=1}^{k} h_i(l) v_i \left(\sum_{s=l}^{k} \frac{m_i(s)}{r_i(s)} \right) - \sum_{l=1}^{k} m_i(l)$$

Subject to $$m_i(l) \geq 0, \forall l$$

Network Problem

$$\mathrm{NET}[\delta; m, \pi, A, c]$$

Maximize $$\sum_{i=1}^{n} \sum_{l=1}^{k} m_i(l) \log(\delta_i(l))$$

Subject to $$\sum_{i \in R_j} \sum_{s=\pi_i(l)}^{k} \delta_i(s) \leq c_j, \forall j, l$$

$$\delta_i(l) \geq 0, \forall i, l$$

- Eisenberg, Gale (1959) - Consensus of subjective probabilities: the pari-mutuel method
- Kelly (1998) - Rate control for communication networks: shadow prices, proportional fairness and stability
- Jain, Vazirani (2010) - Eisenberg-Gale markets: Algorithms and Game Theoretic properties

Equilibrium

Theorem

For any fixed permutation, there exist ***equilibrium***

rates	budgets	incremental allocation	lottery allocations
$r_i^*(l)$	$m_i^*(l)$	$\delta_i^*(l)$	$z_i^*(l)$

such that

(i) $m_i^*(l)$ solves $\mathrm{USER}[m_i; r_i^*, h_i, v_i]$

(ii) $\delta_i^*(l)$ solves $\mathrm{NET}[\delta; m^*, \pi, A, c]$

(iii) $m_i^*(l) = \delta_i^*(l) r_i^*(l)$

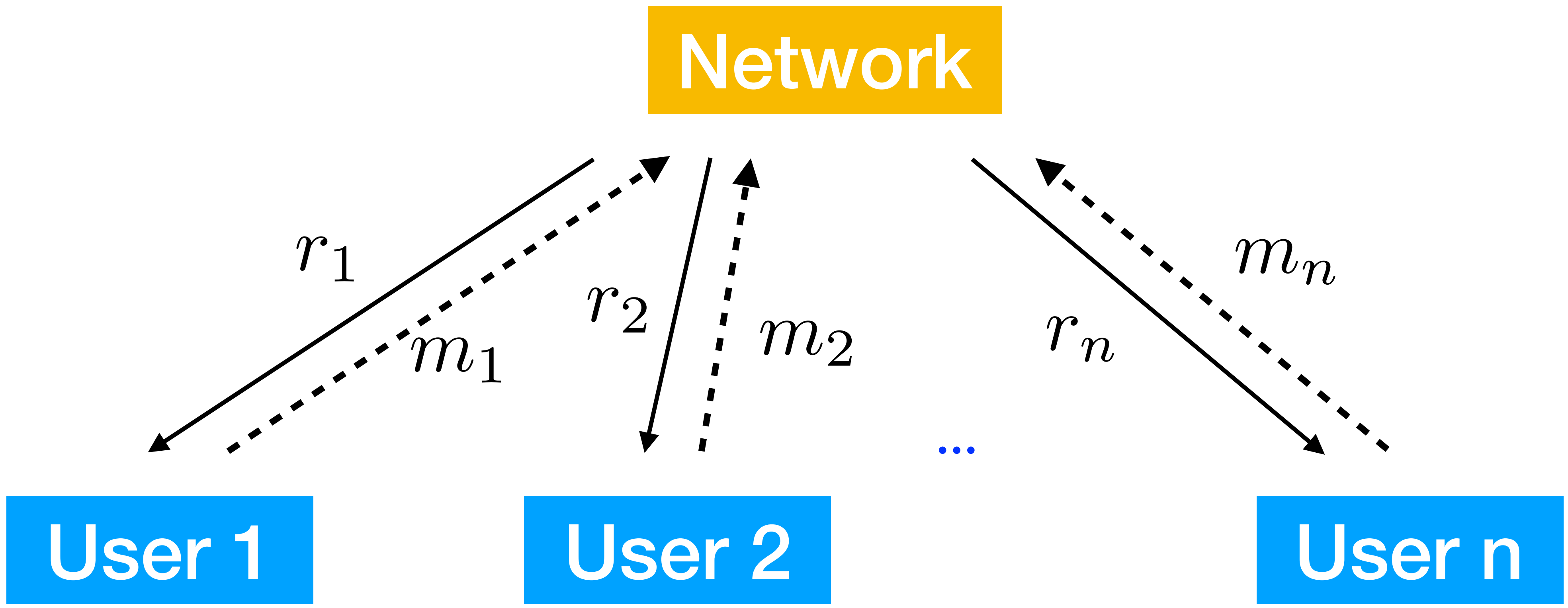
(iv) $\delta_i^*(l) = z_i^*(l) - z_i^*(l+1)$

(v) $z_i^*(l)$ solves the fixed permutation system problem

$$\mathrm{SYS\_FIX}[z; \pi, h, v, A, c]$$

Iterative Process

$\mathrm{NET}[\delta; m, \pi, A, c]$



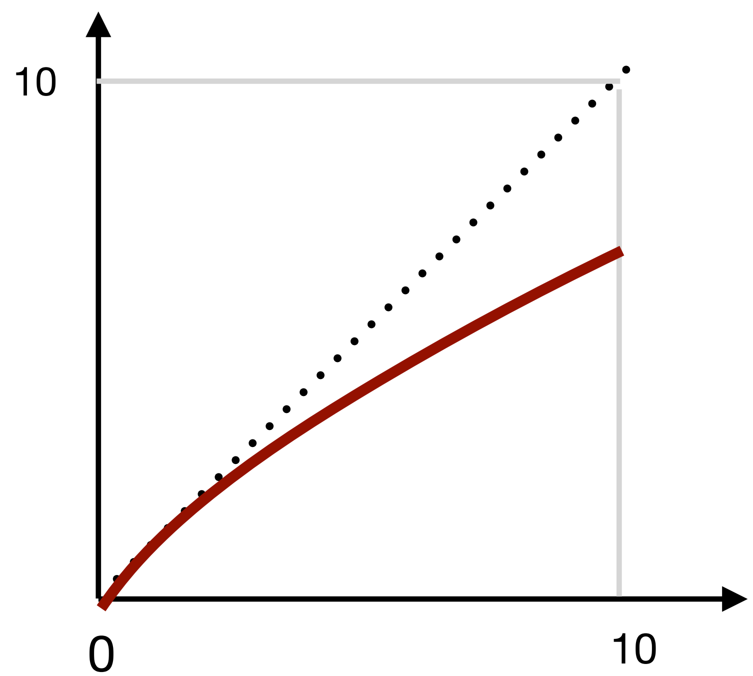
$\mathrm{USER}[m_i; r_i, h_i, v_i]$

Example

n = 10 Players **m = 1 link** **Link capacity c = 10**

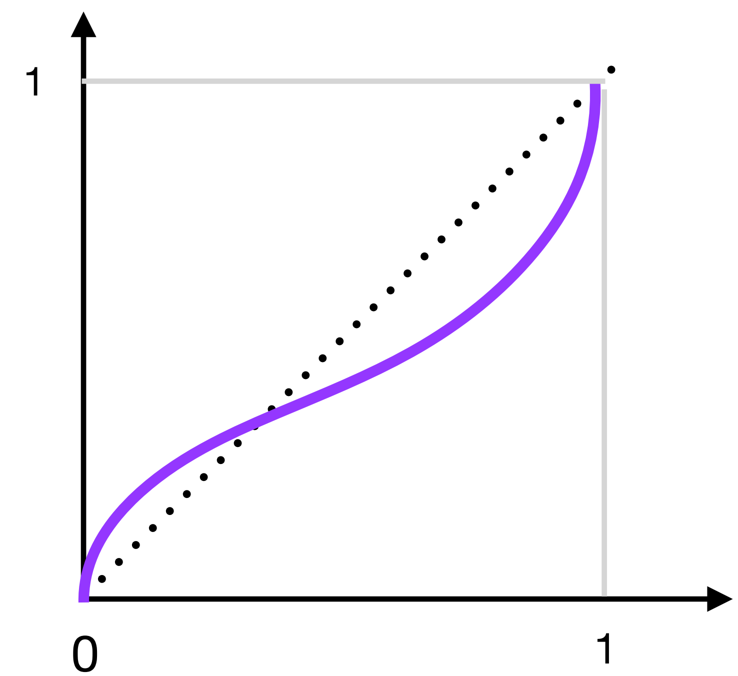
$$v_i(x_i) = x_i^{\beta_i}, \beta_i \in [0,1]$$

$$w_i(p_i) = \frac{p_i^{\gamma_i}}{(p^{\gamma_i} + (1-p)^{\gamma_i})^{1/\gamma_i}}, \gamma_i \in (0,1]$$



Value function

$\beta_i = 0.88$



Probability weighting function

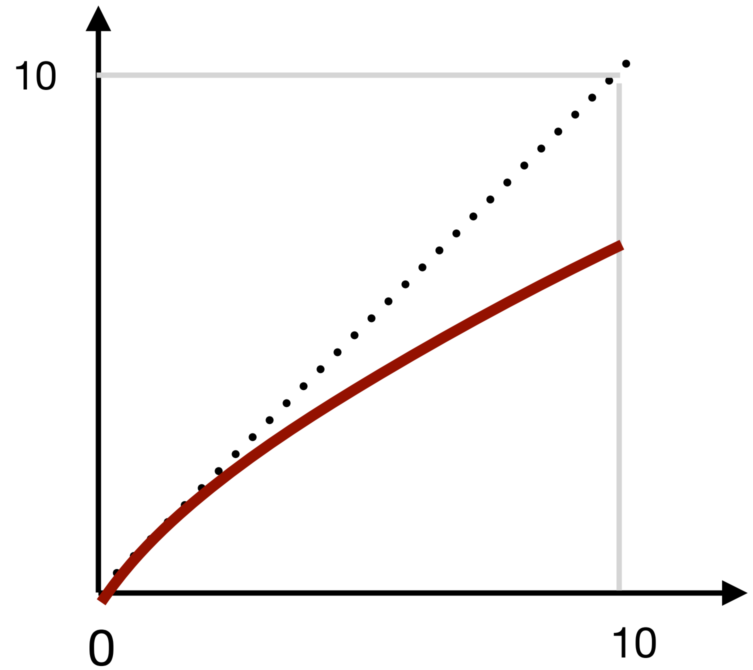
$\gamma_i = 0.61$

Example

n = 10 Players **m = 1 link** **Link capacity c = 10**

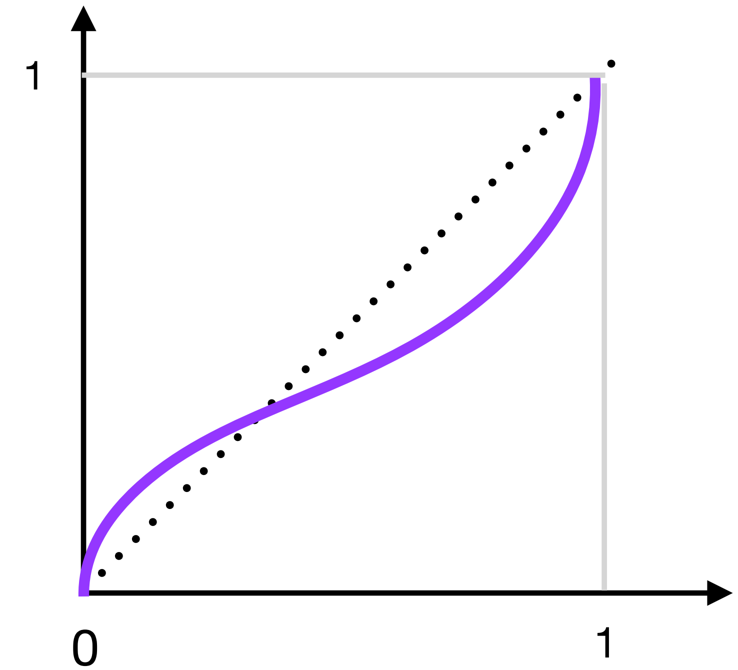
$$v_i(x_i) = x_i^{\beta_i}, \beta_i \in [0,1]$$

$$w_i(p_i) = \frac{p_i^{\gamma_i}}{(p^{\gamma_i} + (1-p)^{\gamma_i})^{1/\gamma_i}}, \gamma_i \in (0,1]$$



Value function

$\beta_i = 0.88$



Probability weighting function

$\gamma_i = 0.61$

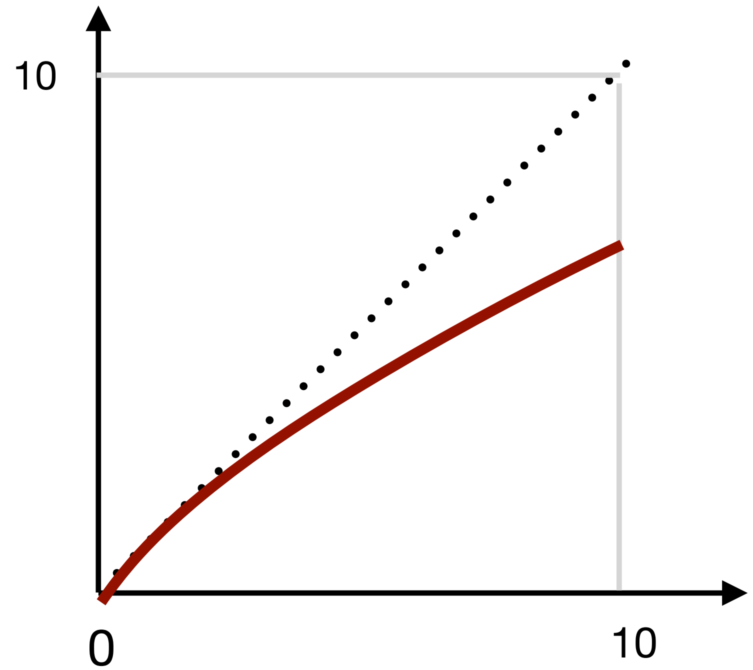
Deterministic Allocation
10

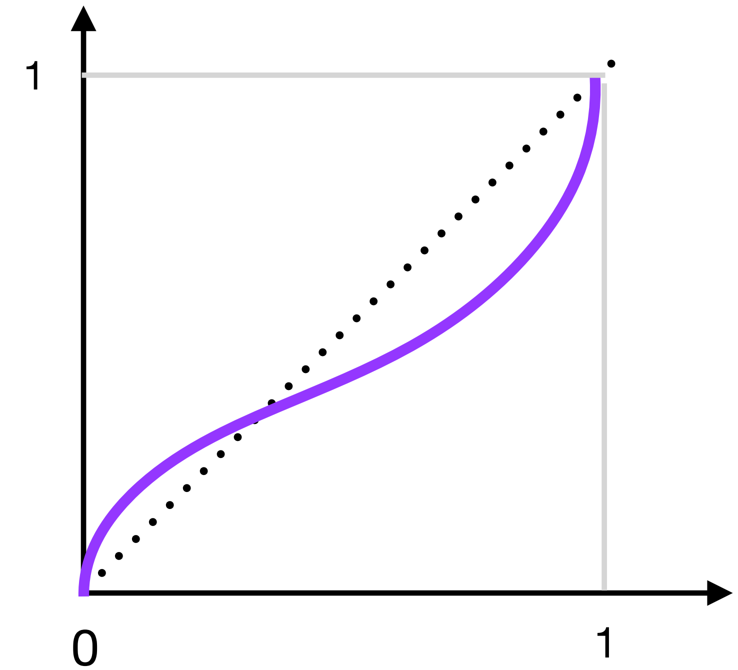
Example

n = 10 Players **m = 1 link** **Link capacity c = 10**

$$v_i(x_i) = x_i^{\beta_i}, \beta_i \in [0,1]$$

$$w_i(p_i) = \frac{p_i^{\gamma_i}}{(p^{\gamma_i} + (1-p)^{\gamma_i})^{1/\gamma_i}}, \gamma_i \in (0,1]$$





Value function

$\beta_i = 0.88$

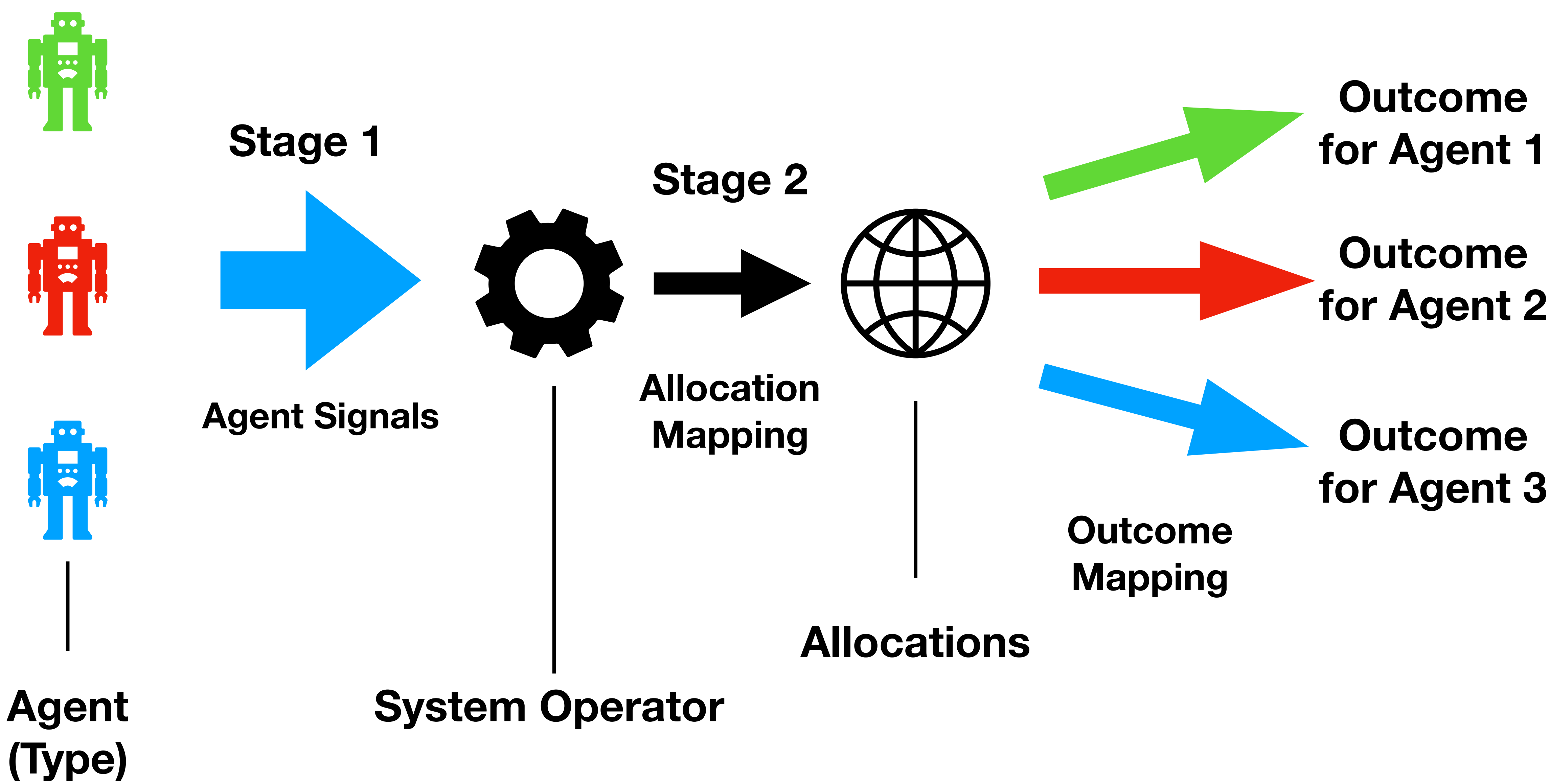
Probability weighting function

$\gamma_i = 0.61$

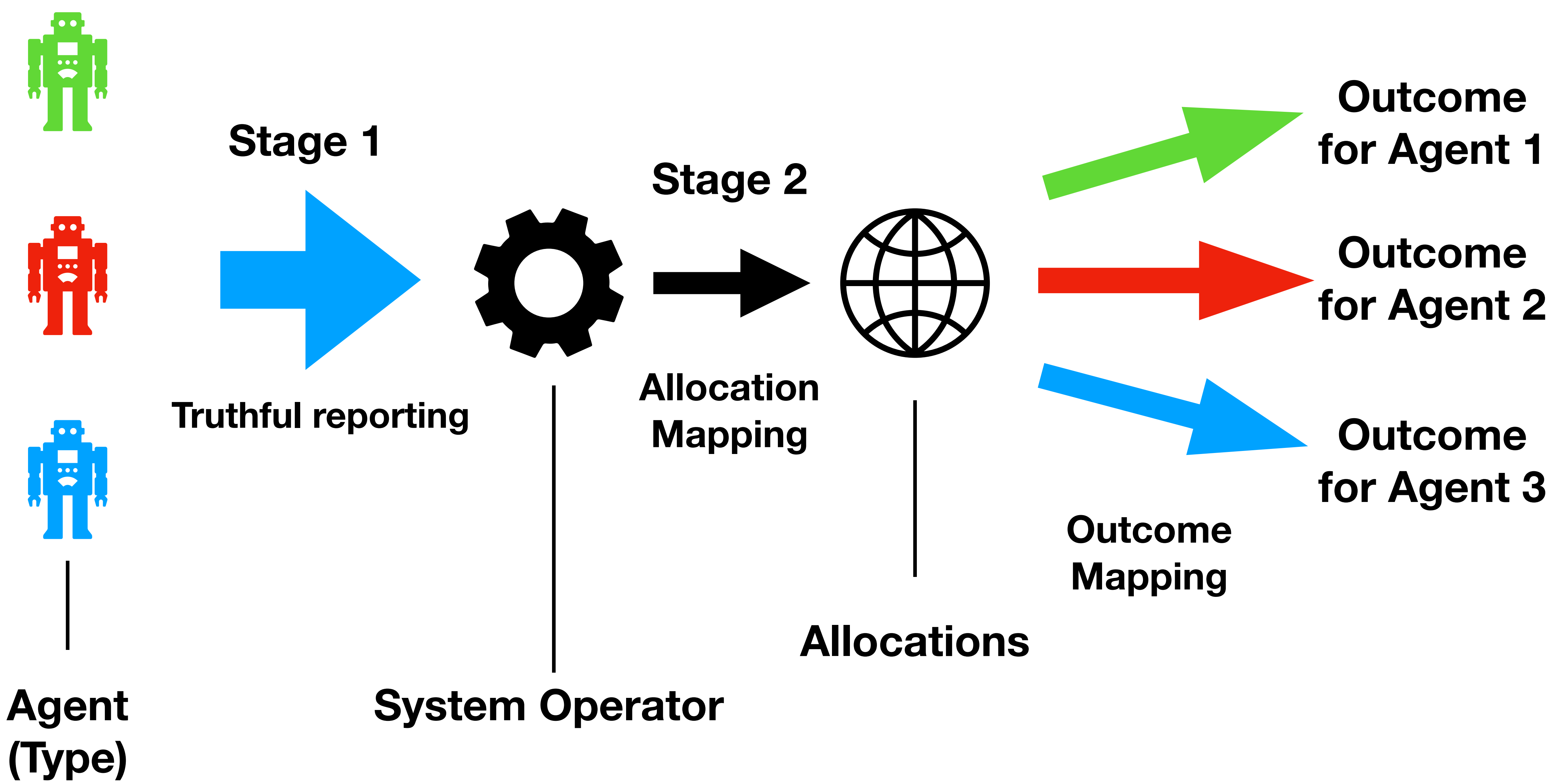
Deterministic Allocation
10

Lottery Allocation
14.17

Mechanism Design



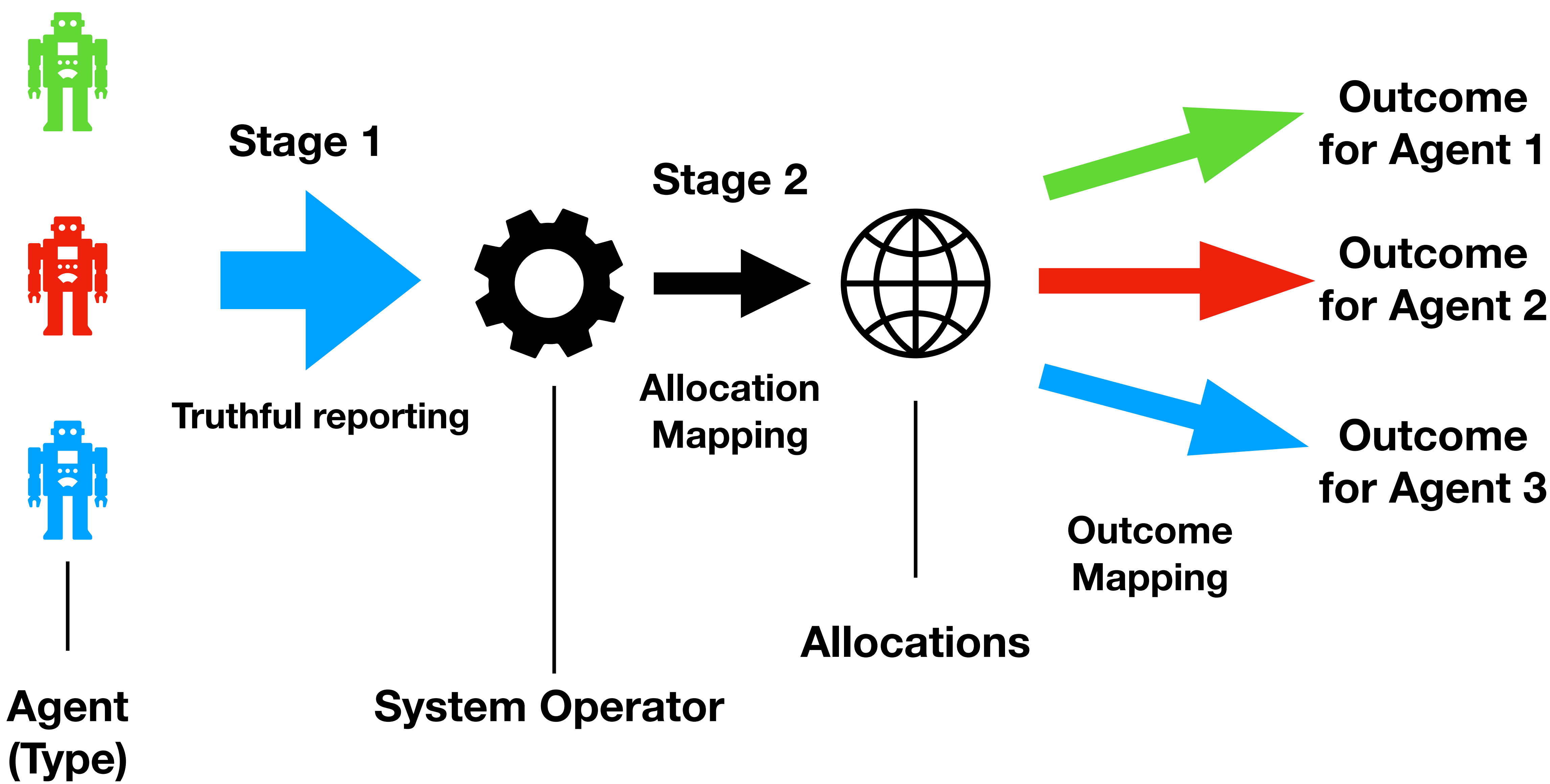
Mechanism Design



Revelation Principle (under EUT)

- WLOG assume signal set = type set for each player
- restrict attention to direct truthful mechanisms

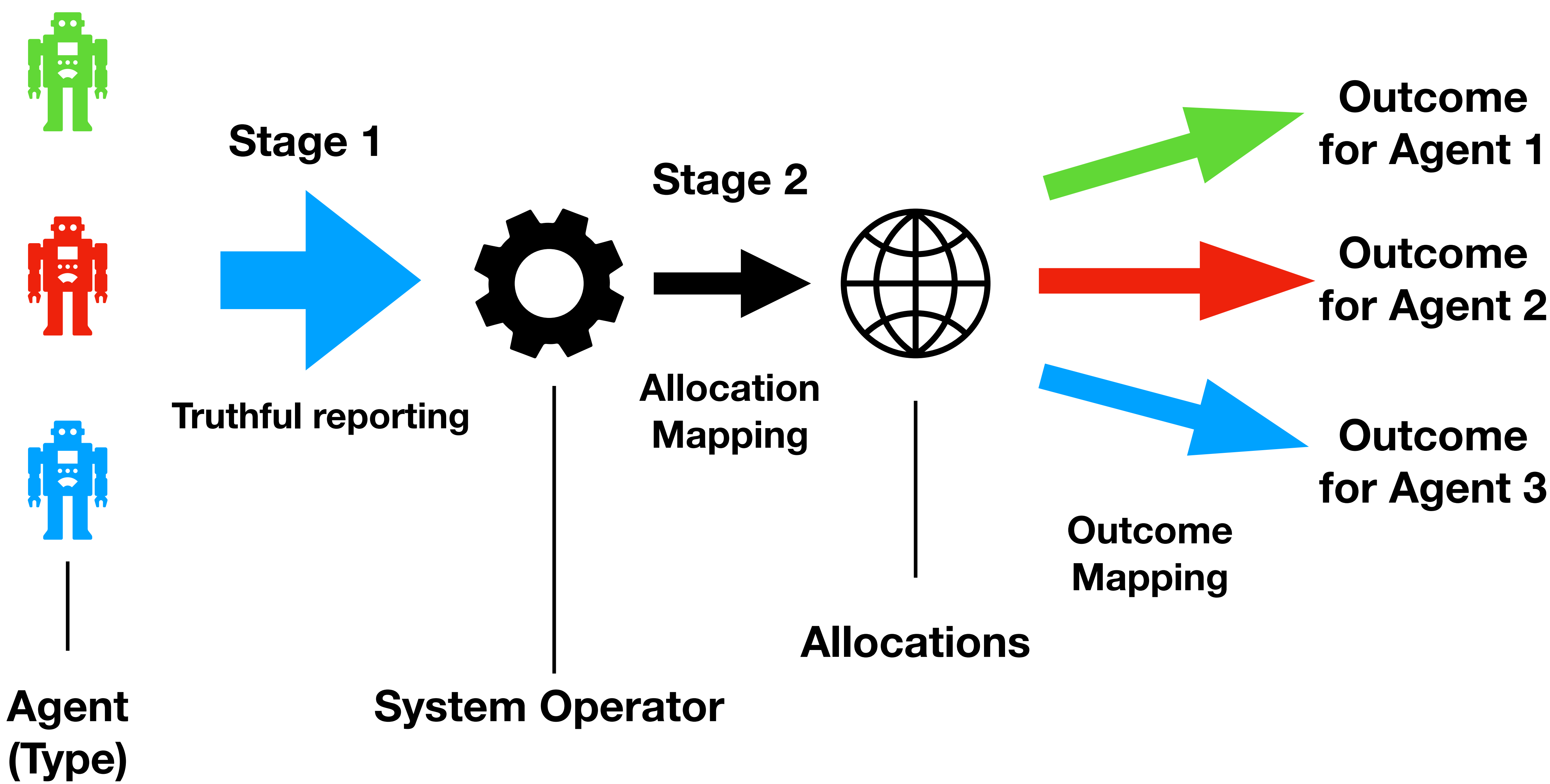
Mechanism Design



Importance of truthful strategies

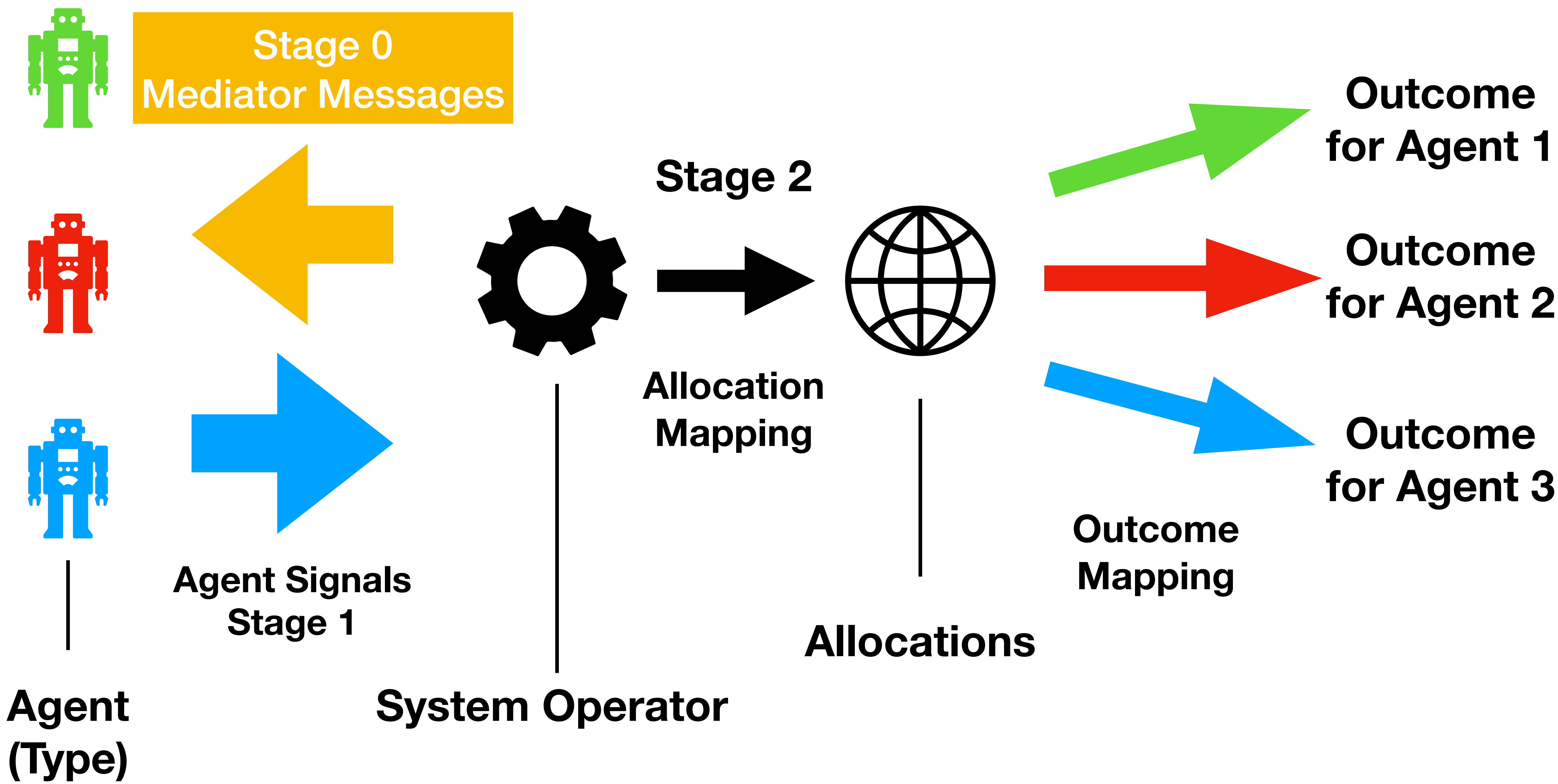
- Limits on information availability
- Computational and cognitive limitations
- Users with different levels of access to information and computation.

Mechanism Design

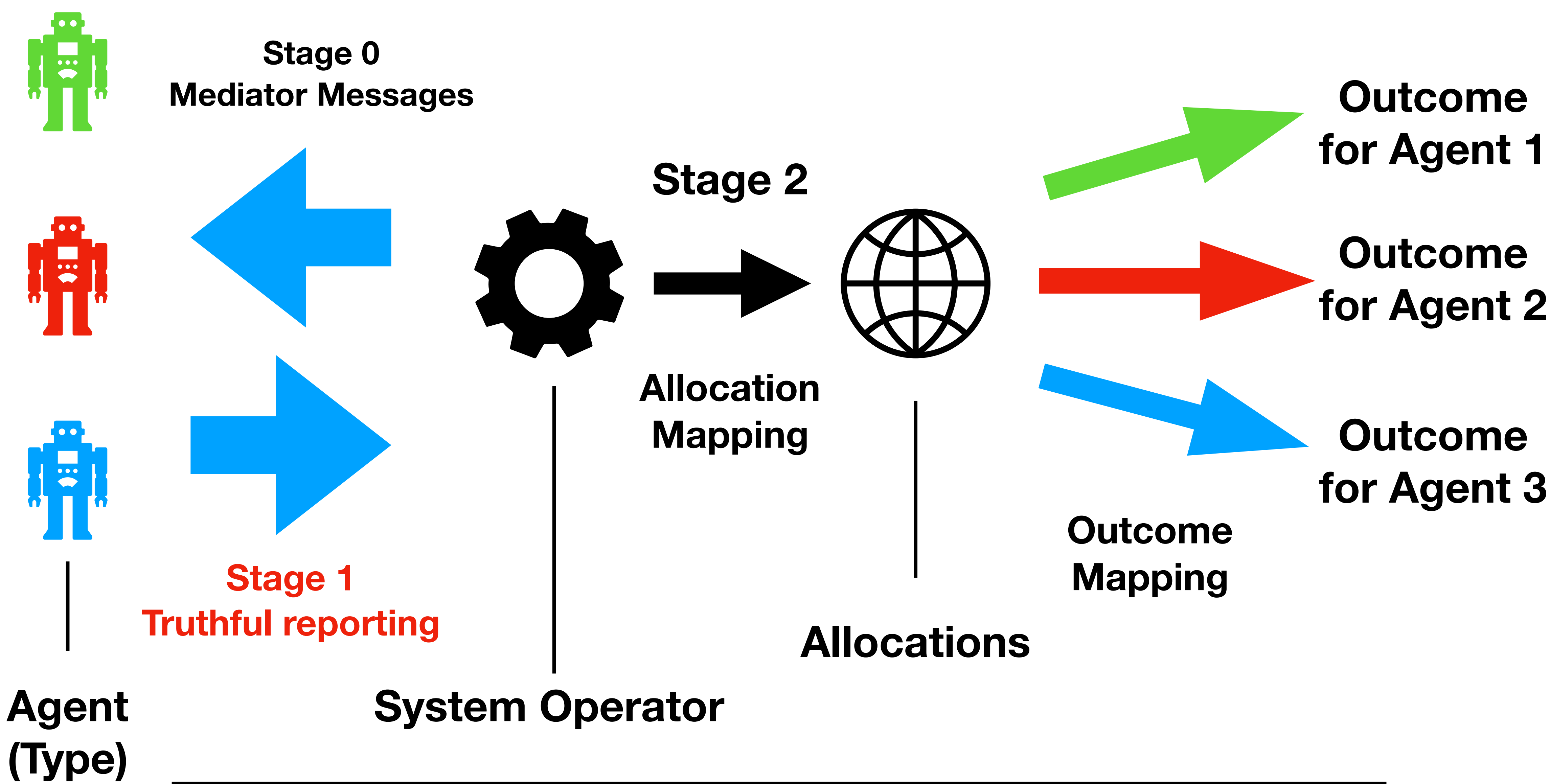


Does not hold under CPT in second-price sealed-bid auctions (Karni and Safra 1989)

Mediated Mechanism Design



Mediated Mechanism Design



Revelation Principle (under CPT)

- WLOG assume signal set = type set for each player
- restrict attention to direct truthful mediated mechanisms

Concluding remarks

- CPT provides a more general framework than EUT.
- CPT seems to more accurately model human agents..
- CPT based designs seem to have tangible benefits.
- Some structural results in EUT continue to hold under CPT in modified form (calibrated learning, mechanism design).
- CPT models provide stronger robustness guarantees relative to the classical techniques of EUT.

Thank you!

No Regret Learning

Question: **Does there exist a learning strategy that does converge to C(Γ)?**

Related to the notion of No regret learning

No Regret Learning

Player 1 imagines replacing his action 0 by action 1

Step	1	2	3	4	5	6
Player 1	**0**	**0**	1	1	1	**0**
Player 2	**I**	**II**	I	III	IV	**I**

L_1 =

Player 2	I	II	III	IV
Outcome	2β	β+1	0	1
Probability	2/3	1/3	0	0

L_2 =

Player 2	I	II	III	IV
Outcome	1.99	1.99	1.99	1.99
Probability	2/3	1/3	0	0

Regret = $(1/2)[V(L_2) - V(L_1)]$

No Regret Learning

Player i imagines replacing his action a_i by $\tilde{a}_i$

$$K_i^t(a_i, \tilde{a}_i) := \xi_i^t(a_i)\Big[V_i\left(\left\{(\xi_{-i}^t(a_{-i}|a_i), \quad x_i(\tilde{a}_i, a_{-i}))\right\}_{l=1}^m\right) - V_i\left(\left\{(\xi_{-i}^t(a_{-i}|a_i), x_i(a_i, a_{-i}))\right\}_{l=1}^m\right)\Big]$$

No Regret Learning

Player i imagines replacing his action a_i by $\tilde{a}_i$

$$K_i^t(a_i, \tilde{a}_i) := \xi_i^t(a_i)\Big[V_i\left(\left\{(\xi_{-i}^t(a_{-i}|a_i), \quad x_i(\tilde{a}_i, a_{-i}))\right\}_{l=1}^m\right) - V_i\left(\left\{(\xi_{-i}^t(a_{-i}|a_i), x_i(a_i, a_{-i}))\right\}_{l=1}^m\right)\Big]$$

- Player i has **no regret learning strategy** if her regrets tend to be arbitrarily small almost surely, irrespective of other players' strategies.

No Regret Learning

Player i imagines replacing his action a_i by $\tilde{a}_i$

$$K_i^t(a_i, \tilde{a}_i) := \xi_i^t(a_i) \Big[V_i \left(\left\{ (\xi_{-i}^t(a_{-i}|a_i), \quad x_i(\tilde{a}_i, a_{-i})) \right\}_{l=1}^m \right) - V_i \left(\left\{ (\xi_{-i}^t(a_{-i}|a_i), x_i(a_i, a_{-i})) \right\}_{l=1}^m \right) \Big]$$

- Player i has **no regret learning strategy** if her regrets tend to be arbitrarily small almost surely, irrespective of other players' strategies.

- No regret learning is equivalent to convergence to empirical distribution

$$\limsup_{t \to \infty} K_i^t(a_i, \tilde{a}_i) \leq 0 \forall a_i, \tilde{a}_i \in A_i \Leftrightarrow \xi^t \to C(\Gamma, i)$$

No Regret Learning

Player i imagines replacing his action a_i by $\tilde{a}_i$

$$K_i^t(a_i, \tilde{a}_i) := \xi_i^t(a_i) \Big[V_i \left(\left\{ (\xi_{-i}^t(a_{-i}|a_i), \quad x_i(\tilde{a}_i, a_{-i})) \right\}_{l=1}^m \right) - V_i \left(\left\{ (\xi_{-i}^t(a_{-i}|a_i), x_i(a_i, a_{-i})) \right\}_{l=1}^m \right) \Big]$$

- Player i has **no regret learning strategy** if her regrets tend to be arbitrarily small almost surely, irrespective of other players' strategies.

- No regret learning is equivalent to convergence to empirical distribution

$$\limsup_{t\to\infty} K_i^t(a_i, \tilde{a}_i) \leq 0 \forall a_i, \tilde{a}_i \in A_i \Leftrightarrow \xi^t \to C(\Gamma, i)$$

- **Question: Does there exists a no regret learning strategy?**

Answer - No! Example

Player 2

Player 1

	I	II	III	IV
0	2β,1	β+1,1	0,1	1,1
1	1.99,0	1.99,0	1.99,0	1.99,0

Action 0 is player 1's Best Response to $\begin{cases} \mu_{odd} = (0.5, 0, 0.5, 0) \\ \mu_{even} = (0, 0.5, 0, 0.5) \end{cases}$

Action 1 is player 1's Best Response to $\mu_{unif} = (0.25, 0.25, 0.25, 0.25)$

Answer - No! Example

Action 0 is player 1's Best Response to $\begin{cases} \mu_{odd} = (0.5, 0, 0.5, 0) \\ \mu_{even} = (0, 0.5, 0, 0.5) \end{cases}$

Action 1 is player 1's Best Response to $\mu_{unif} = (0.25, 0.25, 0.25, 0.25)$

<u>Strategy for player 2</u>

- Play randomized strategy μ_{odd} at step 1,
- Play randomized strategy μ_{even} at step 2,
- Play randomized strategy μ_{odd} at step $2T^k < t \leq T^{k+1}$,
- Play randomized strategy μ_{even} at step $T^{k+1} < t \leq 2T^{k+1}$.

Relaxations

System Prob.

$$\mathrm{SYS}[z, \pi; h, v, A, c]$$

Maximize $\sum_{i=1}^{n} \sum_{l=1}^{k} h_i(l) v_i(z_i(l))$

Subject to $\sum_{i \in R_j} z_i(\pi_i(l)) \leq c_j, \forall j, l$

$z_i(l) \geq z_i(l+1), \forall i, l$

$\pi_i \in S_k, \forall i$

Relaxed System Prob.

$$\mathrm{SYS\_REL}[z, M; h, v, A, c]$$

Maximize $\sum_{i=1}^{n} \sum_{l=1}^{k} h_i(l) v_i(z_i(l))$

Subject to $\sum_{i \in R_j} M_i z_i \leq c_j \mathbf{1}, \forall j$

$z_i(l) \geq z_i(l+1), \forall i, \forall l$

M_i **Doubly Stoc.**

Average System Prob.

$$\mathrm{SYS\_AVG}[z; h, v, A, c]$$

Maximize $\sum_{i=1}^{n} \sum_{l=1}^{k} h_i(l) v_i(z_i(l))$

Subject to $\sum_{i \in R_j} \frac{1}{k} \sum_{l=1}^{k} z_i(l) \leq c_j, \forall j$

$z_i(l) \geq z_i(l+1), \forall i, \forall l$

Theorem: For any system problem,

$$W_{ps} \leq W_{pr} = W_{pa} = W_{da} = W_{dr} = W_{ds}$$

Observations

The Relaxed system problem and the Average system problem are convex optimization problems.

Observations

The Relaxed system problem and the Average system problem are convex optimization problems.

Theorem

The Primal System problem is NP Hard.

Observations

The Relaxed system problem and the Average system problem are convex optimization problems.

Theorem

The Primal System problem is NP Hard.

Proof idea

Integer partition problem can be reduced to a primal system problem and hence NP hard