Game Theory

Lecture 6: Bayesian Games and Mechanism Design

László Gulyás

Eötvös Loránd University, Faculty of Informatics, Department of Artificial Intelligence

2025/26/1

Course textbooks

- Myerson, R. B. (1991). *Game Theory: Analysis of Conflict*. Harvard University Press.
- Nisan, N., Roughgarden, T., Tardos, É., & Vazirani, V. V. (2007). *Algorithmic Game Theory*. Cambridge University Press.
- Shoham, Y., & Leyton-Brown, K. (2008). *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press.
- Bonanno, G. (2024). *Game Theory (3rd ed.)*. University of California, Davis.

Previously on Lecture 5

- **Stochastic games**: state-dependent payoffs and transitions
- **Shapley equation**: fixed point for zero-sum discounted games ($v = Tv$)
- **Value iteration**: converges geometrically to optimal values
- **General-sum** games have Nash equilibria but harder to compute
- **Partial observability**: players maintain beliefs over hidden states
- **Communication**: cheap talk and correlated equilibrium can improve outcomes

What happens when players have private information about their own characteristics?

Lecture Overview

- Real-world motivation: incomplete information in strategic settings
- **Bayesian games**: Type spaces, beliefs, and incomplete information
- **Bayesian Nash equilibrium**: Strategies dependent on private types
- **Perfect Bayesian equilibrium**: Sequential rationality with beliefs
- **Mechanism design**: Reverse game theory - designing games for desired outcomes
- **Revelation principle**: Focus on incentive-compatible mechanisms
- **VCG mechanism**: Efficiency and dominant-strategy incentive compatibility
- **Applications**: Auctions, matching markets, public goods provision

Outline

Incomplete Information in the Real World

Consider these scenarios:

- **Auctions**: Bidders don't know others' valuations for the item
- **Negotiations**: Each side is uncertain about opponent's reservation price
- **Job market**: Employers can't observe worker ability directly
- **Procurement**: Government uncertain about contractor's costs
- **Used car sales**: Buyer doesn't know quality (lemons problem)

*Players have **private information** (types) unknown to others*

From Complete to Incomplete Information

Complete Information	Incomplete Information (Bayesian)
Players know all payoffs	Private types affect payoffs
Nash equilibrium	Bayesian Nash equilibrium
Extensive form (L3)	Bayesian games (L6)
Backward induction	Beliefs and types
Perfect information	Imperfect information about types

Key difference: Players must form **beliefs** about opponents' types

Why “Bayesian”?

Harsanyi’s insight (1967-68): Model incomplete information as:

1. **Nature** draws type profile $\theta = (\theta_1, \ldots, \theta_n)$ from common prior $p(\theta)$
2. Each player i observes only their own type θ_i
3. Players form beliefs about others’ types using **Bayes rule**
4. Play proceeds as game of imperfect information

Transformation: Incomplete information $\rightarrow$ Imperfect information + chance move

Nobel Prize: Harsanyi, Nash, Selten (1994) for equilibrium analysis with incomplete info

Outline

Type Spaces and Beliefs

A **Bayesian game** is:

$$\mathcal{G} = (N, A, \Theta, p, u)$$

- $N = \{1, \ldots, n\}$: Players
- $A = A_1 \times \cdots \times A_n$: Action profiles
- $\Theta = \Theta_1 \times \cdots \times \Theta_n$: Type profiles
- $p(\theta)$: Common prior over types (often independent: $p(\theta) = \prod_i p_i(\theta_i)$)
- $u_i : A \times \Theta \to \mathbb{R}$: Payoffs depend on actions **and** types

Private information: Each player i observes only their own type θ_i

Bayesian Nash Equilibrium

Strategy for player i: $\sigma_i : \Theta_i \to \Delta(A_i)$ (map types to action distributions)

Definition: Strategy profile $\sigma = (\sigma_1, \dots, \sigma_n)$ is a **Bayesian Nash equilibrium (BNE)** if for all i and all $\theta_i \in \Theta_i$:

$$\sigma_i(\theta_i) \in \arg\max_{a_i \in A_i} \mathbb{E}_{\theta_{-i} \sim p(\cdot|\theta_i)}[u_i(a_i, \sigma_{-i}(\theta_{-i}), \theta_i, \theta_{-i})]$$

Intuition: Each type of each player best responds given beliefs about others' types

Special case: When all types are public ($|\Theta_i| = 1$ for all i), BNE = NE

BNE: Key Properties

Existence: Bayesian Nash equilibrium exists (by similar fixed-point argument as Nash)

Computation: Generally harder than NE

- Continuous type spaces $\rightarrow$ infinite-dimensional strategy spaces
- Need to solve for function $\sigma_i(\theta_i)$ rather than finite mixed strategy

Revelation principle: Can focus on direct mechanisms where truth-telling is BNE

Example 1: First-Price Auction

Setup:

- n bidders for single item
- Each bidder i has private valuation $v_i \sim U[0, 1]$ (independent)
- Bidders submit sealed bids $b_i \in [0, \infty)$
- Highest bidder wins, pays their bid
- Ties broken uniformly at random

Payoffs:

$$u_i(b_1, \dots, b_n, v_i) = \begin{cases} v_i - b_i & \text{if } b_i > \max_{j \neq i} b_j \\ 0 & \text{otherwise} \end{cases}$$

Question: What is the Bayesian Nash equilibrium?

First-Price Auction: Solution

Claim: Symmetric BNE has bidding function:

$$b^*(v_i) = \frac{n-1}{n} v_i$$

Verification: If all others bid $b^*(v)$, player i with value v_i solves:

$$\max_{b_i} \Pr[\text{win} \mid b_i] \cdot (v_i - b_i)$$

Given $b_j = \frac{n-1}{n} v_j$ for $j \neq i$, we have $\Pr[\text{win} \mid b_i] = \left(\frac{n}{n-1} b_i\right)^{n-1}$

FOC gives $b_i = \frac{n-1}{n} v_i$ ✓

First-Price Auction: Insights

Shading: Bid below true value (more with more bidders)

Expected revenue:

$$R = \mathbb{E}\left[\max_i b_i\right] = \frac{n-1}{n} \cdot \frac{n}{n+1} = \frac{n-1}{n+1}$$

Winner's curse: Winner often has highest valuation, so be cautious!

Compare to second-price: We'll see later that second-price has $b^*(v_i) = v_i$ (truthful!)

Example 2: Entry Game with Incomplete Information

Setup:

- n potential entrants to a market
- Each has private entry cost $c_i \sim U[0,1]$
- Profit if enter: $\pi(k) = \max\{0, 1 - 0.2k\}$ where k = number of entrants
- Payoff: $\pi(k) - c_i$ if enter, 0 if stay out

Symmetric BNE: Enter if $c_i \leq c^*$ where c^* solves:

$$\pi(\text{expected } \# \text{ entrants}) = c^*$$

Threshold: $c^* \approx 0.5$ for moderate n

Insight: With uncertainty, get probabilistic entry (smoother than all-or-nothing)

Outline

Motivation

Bayesian Games

Perfect Bayesian Equilibrium

Mechanism Design

VCG Mechanism

Applications

Advanced Topics

Summary

Sequential Rationality with Beliefs

Combine **Bayesian games** (types) with **extensive form** (timing):

Perfect Bayesian Equilibrium (PBE) is (σ, μ) where:

1. **Sequential rationality**: At each information set I, strategy $\sigma_i(\cdot \mid I, \theta_i)$ maximizes expected continuation payoff given beliefs μ and opponents' strategies
2. **Belief consistency**: Beliefs μ derived by Bayes rule on path (given σ and prior p)
3. **Off-path beliefs**: Specified to satisfy sequential rationality

Extension of:

- SPE (from L3) $+$ incomplete information
- BNE $+$ sequential structure

Example 3: Entry Deterrence with Incomplete Information

Setup:

- Incumbent is **Tough** (likes fighting, prob p) or **Soft** (dislikes fighting, prob $1-p$)
- Entrant chooses **In** or **Out** (doesn't know incumbent type)
- Incumbent observes entry and chooses **Fight** or **Accommodate**

Payoffs:

	Out	(In, Accommodate)	(In, Fight)
Entrant	0	2	-1
Tough	10	3	**5**
Soft	10	3	**1**

Key: Tough type prefers fighting ($5 > 3$), Soft prefers accommodation ($3 > 1$)

Entry Deterrence: PBE Analysis

Candidate equilibria depend on p:

Case 1: p high $\rightarrow$ Entrant stays **Out** (expects fight)

Case 2: p low $\rightarrow$ Entrant goes **In**, Tough **Fights**, Soft **Accommodates** (separating)

Case 3: p intermediate $\rightarrow$ **Pooling** equilibrium where Soft type mimics Tough!

- Soft fights with some probability to maintain tough reputation
- Entrant mixes between In and Out

Reputation effect: Soft incumbent has incentive to build tough reputation

Example 4: Job Market Signaling (Revisited)

From Lecture 3, now with full PBE analysis:

Sequence:

1. **Nature** draws worker ability $\theta \in \{H, L\}$ with $\pi(H) = 0.3$
2. **Worker** observes θ, chooses education $e \in \{0, 2\}$
3. **Firm** observes e (not θ), offers wage $w \in \{w_L, w_H\}$

Payoffs: Worker gets $w - c(e, \theta)$ where $c(e, H) = e/2$, $c(e, L) = e$ (education costly!)

Question: When can education credibly signal high ability?

Job Market Signaling: Separating PBE

Separating candidate: H chooses $e = 2$, L chooses $e = 0$

Beliefs: $\mu(H \mid e = 2) = 1$, $\mu(H \mid e = 0) = 0$

Firm responses: $w(e = 2) = w_H$, $w(e = 0) = w_L$

Incentive constraints:

- H prefers education: $w_H - 1 \geq w_L$ (cheap for high type)
- L prefers no education: $w_L \geq w_H - 2$ (expensive for low type)

Separating exists if: $w_H - w_L \in [1, 2]$ (education costly enough to separate, not too costly)

Key insight: Even unproductive education can signal if costs differ by type!

Outline

The Reverse Game Theory Problem

Classical game theory:

- **Given**: Game (players, actions, payoffs)
- **Find**: Equilibria and outcomes

Mechanism design:

- **Given**: Desired outcomes (efficiency, fairness, revenue)
- **Design**: Game rules (who moves when, what they know, how payoffs determined)
- **Find**: Mechanisms where equilibrium achieves desired outcome

"Reverse engineering": Choose rules so that rational play implements social goals

Examples of Mechanism Design

Domain	Designer	Goal	Constraints
Auctions	Seller/Auctioneer	Revenue, efficiency	Bidders have private values
Procurement	Buyer	Minimize cost	Suppliers have private costs
Voting	Society	Aggregate preferences	Voters have private types
Matching	Clearinghouse	Stable matches	Agents have private prefs
Public goods	Government	Efficient provision	Citizens have private values

Direct Mechanisms

A **direct mechanism** asks each player to report their type $\widehat{\theta}_i$, then:

1. **Outcome function** $f : \Theta \to X$ determines outcome based on reports
2. **Payment function** $t_i : \Theta \to \mathbb{R}$ determines transfers

Example: In auction, ask for valuations, give item to highest reporter, charge based on reports

Why direct? Revelation principle says we can focus on these WLOG!

Revelation Principle

Theorem (Revelation Principle): Any outcome achievable by any mechanism can be achieved by a **direct mechanism** where:

1. Players truthfully report their types
2. Mechanism computes outcome based on reports

Proof sketch:

- Original mechanism: Players use strategies $s_i(\theta_i)$
- New mechanism: Ask for types, internally compute $s_i(\theta_i)$, run original mechanism
- Truthful reporting is best response (by construction)!

Implication: Focus on **incentive-compatible** direct mechanisms WLOG

Incentive Compatibility

Definition: A direct mechanism (f, t) is **incentive-compatible (IC)** if truth-telling is a (Bayesian Nash) equilibrium:

For all i, all θ_i, all $\hat{\theta}_i$:

$$\mathbb{E}_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i})) - t_i(\theta_i, \theta_{-i})] \geq \mathbb{E}_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i})) - t_i(\hat{\theta}_i, \theta_{-i})]$$

Dominant-strategy IC (DSIC): Truth-telling optimal **regardless** of others' reports (stronger!)

Bayesian IC (BIC): Truth-telling optimal in expectation over others' types

Individual Rationality

Definition: A mechanism is **individually rational (IR)** if each player's expected payoff from participating is non-negative:

$$\mathbb{E}_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i})) - t_i(\theta_i, \theta_{-i})] \geq 0 \quad \forall i, \theta_i$$

Voluntary participation: Players must prefer participating to opting out

Ex-post IR: Non-negative for every type realization (stronger)

Interim IR: Non-negative in expectation over others' types (as above)

Efficiency

Pareto efficiency: Outcome x is PE if no reallocation makes someone better off without hurting others

Ex-post efficiency: For reported types θ, outcome $f(\theta)$ maximizes:

$$\sum_{i=1}^{n} u_i(f(\theta), \theta_i)$$

(utilitarian social welfare)

Goal: Design mechanisms that are IC, IR, **and** efficient

Challenge: Not always possible! (See impossibility theorems)

Outline

Vickrey-Clarke-Groves Mechanism

Setup: Social choice among alternatives X based on private valuations

Mechanism:

1. Each player i reports valuation function $\hat{v}_i : X \to \mathbb{R}$
2. Choose outcome $x^* = \arg\max_{x \in X} \sum_i \hat{v}_i(x)$ (maximize reported total value)
3. Player i pays **Clarke pivot tax**:

$$t_i = \sum_{j \neq i} \hat{v}_j(x^*) - \max_{x \in X} \sum_{j \neq i} \hat{v}_j(x)$$

Intuition: Player pays the **externality** they impose on others

VCG: Dominant Strategy Incentive Compatible

Theorem: In VCG, truth-telling is a **dominant strategy**

Proof: Player i's utility when reporting $\hat{v}_i$:

$$u_i = v_i(x^*) - t_i = v_i(x^*) - \sum_{j \neq i} \hat{v}_j(x^*) + \max_x \sum_{j \neq i} \hat{v}_j(x)$$

Since $x^* = \arg\max_x \sum_j \hat{v}_j(x)$, reporting $\hat{v}_i = v_i$ maximizes:

$$v_i(x^*) + \sum_{j \neq i} \hat{v}_j(x^*)$$

This is **independent** of others' reports (only changes x^* optimally for i) ✓

Key: Payment is independent of own report in the right way!

VCG: Efficiency and Limitations

Efficiency: VCG is **ex-post efficient** (by construction, maximizes social welfare)

Individual rationality: May **violate** IR! (Player may pay more than they value outcome)

Budget balance: Sum of payments may not equal zero (often negative = deficit!)

Vulnerability: Can be manipulated by coalitions (not group-strategy-proof)

Bottom line: VCG achieves DSIC + efficiency, but sacrifices budget balance and sometimes IR

Example 5: Second-Price Auction (Vickrey Auction)

Special case of VCG:

- Alternatives: $X = \{1, \dots, n\}$ (who gets item)
- Valuations: $v_i(x) = v_i$ if $x = i$, else 0

Mechanism:

1. Each bidder reports $\hat{v}_i$
2. Highest bidder wins (allocate efficiently)
3. Winner pays **second-highest bid**

VCG payment:

$$t_i = \max_{j \neq i} \hat{v}_j - \max_j \hat{v}_j = \text{(2nd highest)} - \text{(1st highest)} = -\text{(2nd price)}$$

Result: Bidding true value $b_i = v_i$ is dominant strategy!

Second-Price vs First-Price

Auction Type	Equilibrium Strategy	Revenue	Properties
Second-price	$b^*(v_i) = v_i$ (truthful)	$\mathbb{E}[\text{2nd highest } v]$	DSIC, simple
First-price	$b^*(v_i) = \frac{n-1}{n} v_i$ (shading)	$\frac{n-1}{n+1}$	BIC, strategic

Revenue equivalence theorem: Both auctions yield same expected revenue (in symmetric IPV model)!

Intuition: Shading in first-price exactly compensates for difference in payment rule

Example 6: Public Project

Setup:

- Decide whether to build public project (bridge, park, etc.)
- Cost c to build
- Each player i has private value v_i for project

Efficient outcome: Build if $\sum_i v_i \geq c$

VCG mechanism:

1. Ask for values $\hat{v}_i$
2. Build if $\sum_i \hat{v}_i \geq c$
3. If built, player i pays:

$$t_i = \max\left\{0, c - \sum_{j \neq i} \hat{v}_j\right\}$$

Result: Truth-telling is dominant strategy, project built if and only if efficient!

Outline

Spectrum Auctions

Problem: Allocate radio spectrum licenses to telecom companies

Challenge:

- Companies have private valuations
- Want efficient allocation (licenses to highest-value users)
- Government wants revenue

FCC spectrum auctions:

- Simultaneous ascending auction (related to VCG)
- Raised billions in revenue (e.g., \$19B in 2015 AWS-3 auction)
- Achieved relatively efficient allocation

Innovation: Mechanism design in practice!

Keyword Auctions (Google AdWords)

Setup:

- Advertisers bid for keywords (e.g., "car insurance")
- Multiple ad slots with different click-through rates (CTRs)
- Advertisers have private values-per-click

Generalized Second-Price (GSP) auction:

1. Rank advertisers by bid $\times$ quality score
2. i-th highest bidder gets i-th slot
3. Pays $(i+1)$-th highest bid per click

Not exactly VCG, but approximately incentive-compatible in practice

Revenue: Billions per year for Google!

Kidney Exchange

Problem: Patients need kidney transplants, have willing donors, but blood types incompatible

Solution: Pairwise or chain exchanges

Mechanism design:

- Patients/donors report preferences over potential matches
- Algorithm finds efficient matching (cycles or chains)
- Top Trading Cycles algorithm is strategy-proof!

Real-world impact: Thousands of lives saved annually

Nobel connection: Al Roth (2012 Nobel) pioneered kidney exchange design

Matching Markets

Examples:

- Medical residents to hospitals (NRMP)
- Students to schools
- Workers to firms

Deferred Acceptance (Gale-Shapley):

- Participants report preferences
- Algorithm finds stable matching
- One side's optimal stable matching

Strategy-proofness: Truthful preference revelation is optimal for one side

Used worldwide: NRMP, school choice in NYC/Boston, etc.

Outline

Impossibility Results

Gibbard-Satterthwaite Theorem: Any voting system with ≥ 3 alternatives that is:

- **Onto** (every outcome possible)
- **Deterministic**
- **Strategy-proof** (truth-telling dominant)

must be **dictatorial** (one person always decides)

Implication: Strategic voting unavoidable in most settings!

Budget Balance vs Efficiency

Impossibility (Green-Laffont): Cannot achieve all three of:

1. **Ex-post efficiency**
2. **Budget balance** ($\sum_i t_i = 0$)
3. **Dominant-strategy incentive compatibility**

Trade-offs:

- VCG: Achieves 1 & 3, sacrifices 2
- Mechanism with budget balance: Must sacrifice efficiency or DSIC

Optimal Auctions

Myerson's optimal auction: Maximize seller revenue subject to IC and IR

Result:

- Set reserve price $r > 0$ (exclude low-value bidders)
- Use virtual valuations: $\varphi_i(v_i) = v_i - \frac{1-F_i(v_i)}{f_i(v_i)}$
- Allocate to bidder with highest positive virtual valuation

Key insight: Revenue-maximizing mechanism may sacrifice efficiency!

When to Use Bayesian Game Theory

Use Bayesian games when:

- **Private information** critical to strategic interaction
- **Beliefs** about others' types matter
- **Screening** or **signaling** important

Examples:

- Auctions (private valuations)
- Negotiations (reservation prices unknown)
- Job market (worker ability private)
- Insurance (risk types private)

When to Use Mechanism Design

Use mechanism design when:

- You **control the rules** of interaction
- Want to **implement** specific outcomes (efficiency, fairness, revenue)
- Need to **elicit private information**

Examples:

- Running an auction (seller)
- Allocating resources (government, firm)
- Matching markets (clearinghouse)
- Voting systems (social planner)

Key Takeaways

1. **Bayesian games** model incomplete information via type spaces and beliefs
2. **Bayesian Nash equilibrium**: Strategies depend on beliefs over types
3. **Perfect Bayesian equilibrium**: Sequential rationality + belief consistency
4. **Mechanism design**: Reverse engineer games to achieve desired outcomes
5. **Revelation principle**: Focus on truthful direct mechanisms WLOG
6. **VCG mechanism**: Efficient and DSIC, but not budget-balanced
7. **Trade-offs**: Cannot always achieve efficiency + budget balance + IC + IR
8. **Applications**: Auctions, matching markets, voting, public goods

Questions to Consider

1. How do beliefs about others' types affect equilibrium strategies?
2. When is separating equilibrium possible in signaling games?
3. Why is second-price auction truthful but first-price is not?
4. What trade-offs exist between efficiency and budget balance?
5. How can mechanism design improve real-world allocation problems?

Exercises

Exercise 1: In first-price auction with 3 bidders and $v_i \sim U[0,1]$, verify that $b^*(v_i) = \frac{2}{3}v_i$ is BNE.

Exercise 2: In entry game with incomplete information, find the symmetric BNE threshold c^* when $n = 5$ and $\pi(k) = 1 - 0.2k$.

Exercise 3: Show that in second-price auction, bidding true value $b_i = v_i$ is a weakly dominant strategy.

Exercise 4: Design a VCG mechanism for allocating a public good. Show truth-telling is DSIC.

Exercise 5: Prove that if a mechanism is DSIC and efficient, it cannot be budget-balanced (for generic valuations).

Outline

Summary

Today we covered:

- Bayesian games: type spaces, beliefs, incomplete information
- Bayesian Nash equilibrium and Perfect Bayesian equilibrium
- Mechanism design: reverse game theory
- Revelation principle and incentive compatibility
- VCG mechanism: efficiency and DSIC
- Applications: auctions, matching, public goods
- Trade-offs and impossibility results

Thank You

Thank You!