Game Theory 2025

Ulle Endriss
Institute for Logic, Language and Computation
University of Amsterdam

Plan for Today

In this second lecture on *mechanism design* we are going to generalise beyond the basic scenario of auctions—as much as we can manage:

- *Revelation Principle:* can focus on direct-revelation mechanisms
- formal model of *direct-revelation mechanisms with money*
- *incentive compatibility* of the *Vickrey-Clarke-Groves mechanism*
- other *properties* of VCG for special case of *combinatorial auctions*
- *impossibility* of achieving incentive compatibility more generally

Much of this is also (somewhat differently) covered by Nisan (2007).

N. Nisan. Introduction to Mechanism Design (for Computer Scientists). In N. Nisan et al. (eds.), *Algorithmic Game Theory*. Cambridge University Press, 2007.

Reminder

Last time we saw *four auction mechanisms* for selling a single item: English, Dutch, first-price sealed-bid, Vickrey.

The *Vickrey auction* was particularly interesting:

- each bidder submits a bid in a sealed envelope
- the bidder with the highest bid wins, but pays the price of the second highest bid (unless it's below the reservation price)

It is a *direct-revelation mechanism* (unlike English and Dutch auctions) and it is *incentive-compatible*, i.e., truth-telling is a dominant strategy (unlike Dutch and FPSB auctions).

The Revelation Principle

Revelation Principle: Any outcome that is implementable in dominant strategies via *some* mechanism can also be implemented by means of a *direct-revelation mechanism* making *truth-telling* a *dominant strategy*.

This can be formulated (and proved) as a formal theorem, but here we are going to be content with understanding the underlying intuition:

> *Given mechanism M, build a corresponding direct-revelation mechanism M' as follows. Ask each player i for a valuation $\hat{v}_i$. Then simulate M being played by rational agents whose true valuations are $(\hat{v}_1, \ldots, \hat{v}_n)$. This makes submitting your true valuation a dominant strategy, as that will enable the agent playing inside the black box on your behalf to do so optimally.*

Thus: Sufficient to focus on direct-revelation mechanisms from now on.

Example: Can think of the *Vickrey auction* as direct-revelation variant of the *English auction* (with arbitrarily low increments ϵ).

Direct-Revelation Mechanisms with Money

A *direct-revelation mechanism* is a tuple $\langle N, \Omega, \boldsymbol{V}, f, \boldsymbol{p} \rangle$, where:

- $N = \{1, \ldots, n\}$ is a finite set of *agents*,
- $\Omega = \{1, \ldots, m\}$ is a finite set of possible *outcomes*,
- $\boldsymbol{V} = V_1 \times \cdots \times V_n$, with sets $V_i \subseteq \mathbb{R}^\Omega$ of *possible valuations* for i,
- $f : \boldsymbol{V} \to \Omega$ is a *social choice function*, and
- $\boldsymbol{p} = (p_1, \ldots, p_n)$ is a profile of *price functions* $p_i : \boldsymbol{V} \to \mathbb{R}$.

Each agent i has a (private) *true valuation* $v_i : \Omega \to \mathbb{R}$ with $v_i \in V_i$.

Each agent i submits a *bid* $\hat{v}_i : \Omega \to \mathbb{R}$ with $\hat{v}_i \in V_i$, which may or may not be equal to v_i, resulting in a profile $\hat{\boldsymbol{v}} = (\hat{v}_1, \ldots, \hat{v}_n)$.

Then outcome $f(\hat{\boldsymbol{v}})$ gets implemented, with prices $(p_1(\hat{\boldsymbol{v}}), \ldots, p_n(\hat{\boldsymbol{v}}))$, and agent i experiences (quasi-linear) *utility* $u_i(\hat{\boldsymbol{v}}) = v_i(f(\hat{\boldsymbol{v}})) - p_i(\hat{\boldsymbol{v}})$.

Remark: Observe how *combinatorial auctions*, with Ω being the set of all possible allocations of goods to bidders, are a special case.

Connection to Bayesian Games

This model is similar to that of *Bayesian games* $\langle N, \boldsymbol{A}, \boldsymbol{\Theta}, p, \boldsymbol{u} \rangle$ ...

- N: players are agents
- $\boldsymbol{A}$: actions are declared valuations
- $\boldsymbol{\Theta}$: types are true valuations
- $\boldsymbol{u}$: utilities are determined by the true valuations, together with the social choice function and the price functions

Only common prior p is missing. So agents only know what valuations others might have, not how probable any given situation is.

That's ok: we want to study dominant strategies only, i.e., strategies that work best for *all* type profiles that are possible at all.

Incentive Compatibility

Our main property of interest for today is incentive compatibility . . .

A mechanism $\langle N, \Omega, \boldsymbol{V}, f, \boldsymbol{p}\rangle$ is called *incentive-compatible* in case *truth-telling* is a *dominant strategy* for every agent $i \in N$:

$$v_i(f(v_i, \hat{\boldsymbol{v}}_{-i})) - p_i(v_i, \hat{\boldsymbol{v}}_{-i}) \quad \geqslant \quad v_i(f(\hat{v}_i, \hat{\boldsymbol{v}}_{-i})) - p_i(\hat{v}_i, \hat{\boldsymbol{v}}_{-i})$$

for all $v_i, \hat{v}_i \in V_i$ and $\hat{\boldsymbol{v}}_{-i} \in \boldsymbol{V}_{-i}$

An alternative term for this concept is *strategyproofness*.

Recall: The Vickrey auction is incentive-compatible.

Formally Modelling the Vickrey Auction

The Vickrey auction is the mechanism $\langle N, \Omega, \boldsymbol{V}, f, \boldsymbol{p} \rangle$, where:

- $N = \{1, \ldots, n\}$ is the set of bidders;
- $\Omega = N$, with outcome $i \in \Omega$ expressing that bidder $i \in N$ wins;
- $\boldsymbol{V} = V_1 \times \cdots \times V_n$, with $V_i = \{\hat{v}_i : x \mapsto w \cdot \mathbb{1}_{x=i} \mid w \in \mathbb{R}_{\geqslant 0}\}$ and $\hat{v}_i(x)$ denoting the valuation (potentially) declared by bidder i for the outcome under which player x receives the item;
- $f : (\hat{v}_1, \ldots, \hat{v}_n) \mapsto \mathrm{argmax}_i\, \hat{v}_i(i)$ selects the highest bid; and
- $\boldsymbol{p} = (p_1, \ldots, p_n)$, with price functions p_i defined as follows:

$$p_i(\hat{\boldsymbol{v}}) \;=\; \begin{cases} \max\{\, \hat{v}_j(j) \mid j \in N \setminus \{i)\}\,\} & \text{if } i = f(\hat{\boldsymbol{v}}) \\ 0 & \text{otherwise} \end{cases}$$

In practice, *tie-breaking* between maximal bids needs to be dealt with.

<u>Exercise:</u> *Can you adapt Vickrey's idea to combinatorial auctions?*

Generalisation to Combinatorial Auctions

Recall: In a CA, each bidder i has a valuation $v_i : 2^G \to \mathbb{R}_{\geqslant 0}$ on bundles of goods and she bids by reporting some $\hat{v}_i : 2^G \to \mathbb{R}_{\geqslant 0}$.

The set of *outcomes* Ω is the set of all *allocations* of the form $(B_1, \ldots, B_n) \in 2^G \times \cdots \times 2^G$ with $B_i \cap B_j = \emptyset$ for all $i, j \in N$.

Now think of valuations being applied to *allocations* $\omega = (B_1, \ldots, B_n)$, rather than to *bundles* B_i: thus, write $v_i(\omega)$ for $v_i(B_i)$.

This encoding allows us to abstract away from CAs and to carry out the following analysis for arbitrary direct-revelation mechanisms . . .

Alternative Interpretation of Vickrey's Pricing Rule

Idea: In a Vickrey auction, the winner pays her bid, but gets a *discount*.

How much? The size of the discount reflects the *marginal contribution* to social welfare made by the winner:

- Without the winner's bid, the second highest bid would have won. So the marginal contribution made by the winner is equal to the difference between the winning and the second highest bid.
- Subtracting this marginal contribution from the winning bid yields the second highest bid (the Vickrey price).

Generalisation of Vickrey's Idea

The *social choice function* $f : \boldsymbol{V} \to \Omega$ maps every profile of reported valuations to an outcome maximising (reported) social welfare:

$$f(\hat{\boldsymbol{v}}) \in \operatorname*{argmax}_{\omega \in \Omega} \sum_{i \in N} \hat{v}_i(\omega)$$

The *price function* $p_i : \boldsymbol{V} \to \mathbb{R}$ charges agent i the price she offered for what gets implemented, minus the marginal contribution she made:

$$\begin{aligned} p_i(\hat{\boldsymbol{v}}) &= \hat{v}_i(f(\hat{\boldsymbol{v}})) - \left(\sum_{j=1}^{n} \hat{v}_j(f(\hat{\boldsymbol{v}})) - \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}}_{-i})) \right) \\ &= \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}}_{-i})) - \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}})) \end{aligned}$$

Thus: Agent i pays sum of the losses in value she causes for others.

Remark: So f should be defined for $n-1$ agents as well.

The Vickrey-Clarke-Groves Mechanism

What we just defined is the famous *VCG mechanism*. Good news:

Theorem 1 *The VCG mechanism is incentive-compatible.*

Proof: Consider agent i. She pays $p_i(\hat{\boldsymbol{v}}) = h_i(\hat{\boldsymbol{v}}_{-i}) - \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}}))$, where $h_i(\hat{\boldsymbol{v}}_{\boldsymbol{-i}})$ is a term she can't affect. Her utility is $v_i(f(\hat{\boldsymbol{v}})) - p_i(\hat{\boldsymbol{v}})$.

So she *wants* to maximise the term $v_i(f(\hat{\boldsymbol{v}})) + \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}}))$, while the mechanism actually maximises $\hat{v}_i(f(\hat{\boldsymbol{v}})) + \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}}))$.

Thus, the best she can do is to report $\hat{v}_i = v_i$. ✓

Exercise: *Do you see the immediate generalisation suggesting itself?*

W. Vickrey. Counterspeculation, Auctions, and Competitive Sealed Tenders. *Journal of Finance*, 16(1):8–37, 1961.

E.H. Clarke. Multipart Pricing of Public Goods. *Public Choice*, 11(1):17–33, 1971.

T. Groves. Incentives in Teams. *Econometrica*, 41(4):617–631, 1973.

The Family of Groves Mechanisms

A *Groves mechanism* is a mechanism where f maximises social welfare and price functions p_i are of this form, for some function h_i:

$$p_i(\hat{\boldsymbol{v}}) \quad = \quad h_i(\hat{\boldsymbol{v}}_{\boldsymbol{-i}}) - \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}}))$$

Theorem 2 (Groves, 1973) *Any mechanism belonging to the family of Groves mechanisms is incentive-compatible.*

The proof is the same as before. *VCG* is the Groves mechanism with the so-called *Clarke tax* $h_i(\hat{\boldsymbol{v}}_{-i}) = \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}}_{-i}))$.

Remark: By the *Green-Laffont Theorem*, Groves mechanisms are also *the only* mechanisms that maximise SW and are incentive-compatible.

T. Groves. Incentives in Teams. *Econometrica*, 41(4):617–631, 1973.

J. Green and J.-J. Laffont. Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods. *Econometrica*, 45(2): 427–438, 1977.

Weak Budget Balance

Why choose the Clarke tax? Maximal social welfare does *not* imply maximal revenue for the auctioneer. But she should not lose money ...

A mechanism is *weakly budget balanced* if it guarantees $\sum_i p_i(\hat{\boldsymbol{v}}) \geqslant 0$.

Proposition 3 *Assuming free disposal, the VCG mechanism for combinatorial auctions is weakly budget balanced.*

Here *free disposal* means that the auctioneer need not sell all goods.

Proof: We prove $p_i(\hat{\boldsymbol{v}}) \geqslant 0$ for all $i \in N$, which is stronger.

The following holds due to free disposal. It expresses that we can get higher SW for $N \setminus \{i\}$ if we optimise for $N \setminus \{i\}$ rather than for N:

$$\sum_{j \in N \setminus \{i\}} \hat{v}_j(f(\hat{\boldsymbol{v}}_{-\boldsymbol{i}})) \;\; \geqslant \;\; \sum_{j \in N} \hat{v}_j(f(\hat{\boldsymbol{v}})) \; - \; \hat{v}_i(f(\hat{\boldsymbol{v}}))$$

Thus: $p_i(\hat{\boldsymbol{v}}) = \hat{v}_i(f(\hat{\boldsymbol{v}})) - \left[\sum_j \hat{v}_j(f(\hat{\boldsymbol{v}})) - \sum_{j \neq i} \hat{v}_j(f(\hat{\boldsymbol{v}}_{-i}))\right] \geqslant 0.$ ✓

Remark: Also works for *monotonic valuations* (instead of free disposal).

Issues with VCG

So: VCG maximises social welfare (though not revenue), and it is both weakly budget balanced and incentive-compatible. Nice.

But VCG is not perfect. Some concerns (for *combinatorial auctions*):

- *Low revenue* for the auctioneer, possibly even *zero revenue*
- *Failure of monotonicity:* additional bids may decrease revenue
- *Collusion:* incentive-compatibility does not protect from coalitions of bidders coordinating their untruthful bids and sharing the profit
- *False-name bidding:* a bidder may benefit from bidding separately under multiple identities

We are now going to see examples for a couple of these.

L.M. Asubel and P. Milgrom. The Lovely but Lonely Vickrey Auction. In P. Cramton et al. (eds.), *Combinatorial Auctions*. MIT Press, 2006.

Example: Zero Revenue

There are cases where the VCG mechanism generates zero revenue.

Suppose there are just two goods and bidders list prices for all bundles.

Bidder 1: $(\emptyset, 0)$, $(\{a\}, 0)$, $(\{b\}, 0)$, $(\{a, b\}, 2)$

Bidder 2: $(\emptyset, 0)$, $(\{a\}, 2)$, $(\{b\}, 0)$, $(\{a, b\}, 0)$

Bidder 3: $(\emptyset, 0)$, $(\{a\}, 0)$, $(\{b\}, 2)$, $(\{a, b\}, 0)$

Payments are computed as follows:

Bidder 1: 0

Bidder 2: $2 - (4 - 2) = 0$

Bidder 3: $2 - (4 - 2) = 0$

Example: False-Name Bidding

Suppose there are two bidders and two goods:

Bidder 1: $(\emptyset, 0)$, $(\{a\}, 0)$, $(\{b\}, 0)$, $(\{a, b\}, 4)$

Bidder 2: $(\emptyset, 0)$, $(\{a\}, 1)$, $(\{b\}, 1)$, $(\{a, b\}, 2)$

Bidder 1 wins. But bidder 2 can instead submit bids *under two names:*

Bidder 1: $(\emptyset, 0)$, $(\{a\}, 0)$, $(\{b\}, 0)$, $(\{a, b\}, 4)$

Bidder 2: $(\emptyset, 0)$, $(\{a\}, 4)$, $(\{b\}, 0)$, $(\{a, b\}, 0)$

Bidder 2′: $(\emptyset, 0)$, $(\{a\}, 0)$, $(\{b\}, 4)$, $(\{a, b\}, 0)$

Bidder(s) 2 (and 2′) will win and not pay anything! This form of manipulation is particularly critical for *electronic* auctions, as it is easier to create multiple identities online than it is in real life.

M. Yokoo. Pseudonymous Bidding in Combinatorial Auctions. In P. Cramton et al. (eds.), *Combinatorial Auctions*, MIT Press, 2006.

Mechanism Design without Money

We've generalised from single-item auctions, to combinatorial auctions, to arbitrary direct-revelation mechanisms with money. *What more?*

Bidders have specific kinds of *preferences* over outcomes-plus-prices. We also might attempt mechanism design for arbitrary preferences: *weak orders* over outcomes (w/o separating prices).

Recall: weak order = binary relation that is transitive + complete

Now a *mechanism* (w/o money) is a tuple $\langle N, \Omega, f \rangle$, where:

- $N = \{1, \ldots, n\}$ is a finite set of *agents*,
- $\Omega = \{1, \ldots, m\}$ is a set of *outcomes*, and
- $f : (2^{\Omega \times \Omega})^n \to \Omega$ is a *social choice function*, mapping any given profile of reported weak orders on Ω to an outcome in Ω.

It is (tedious but) possible to translate mechanisms with money into mechanisms w/o money. Exercise: *Sketch how to do this!*

The Gibbard-Satterthwaite Theorem

Now *incentive-compatibility* means: $f(\succeq_i, \hat{\succeq}_{-i}) \succeq_i f(\hat{\succeq}_i, \hat{\succeq}_{-i})$.

Theorem 4 (Gibbard-Satterthwaite) *A mechanism $\langle N, \Omega, f\rangle$ with $|\Omega| \geqslant 3$ and surjective f is incentive-compatible <u>iff</u> it is a dictatorship.*

Here a *dictatorship* always chooses the top outcome of the same player.

For an accessible proof, consult my expository paper cited below.

<u>Remark:</u> We can think of f as a *voting rule* (Ω are the candidates).

A. Gibbard. Manipulation of Voting Schemes: A General Result. *Econometrica*, 41(4):587–601, 1973.

M.A. Satterthwaite. Strategy-proofness and Arrow's Conditions. *Journal of Economic Theory*, 10:187–217, 1975.

U. Endriss. Logic and Social Choice Theory. In A. Gupta and J. van Benthem (eds.), *Logic and Philosophy Today*, College Publications, 2011.

Summary

This has been the second and final lecture on *mechanism design*.

Our main objective has been to find *incentive-compatible* mechanisms. The central idea in the work of Vickrey, Clarke, and Groves to achieve this is to maximise *social welfare* and charge *prices* like this:

- you pay what you offered, but get a discount equal to the increase in (*total*) social welfare caused by your participation
- (alternative reading) you pay the loss in social welfare experienced *by the others* caused by your participation

Some generalisation (Groves mechanisms) is possible, but there are limitations (*Green-Laffont* and *Gibbard-Satterthwaite* Theorems).

What next? We switch to cooperative game theory and study how players form coalitions and divide the value they produce together.