Obviously Strategyproof Mechanisms

ETH Zürich Paolo Penna

In the previous lecture we have seen that search engines use the Generalized Second Price mechanism despite it is not truthful. One of the advantages of GSP is perhaps that it is simpler for the players to understand, compared to the truthful VCG mechanism. Consider the two equivalent ways to run a truthful single item auction:

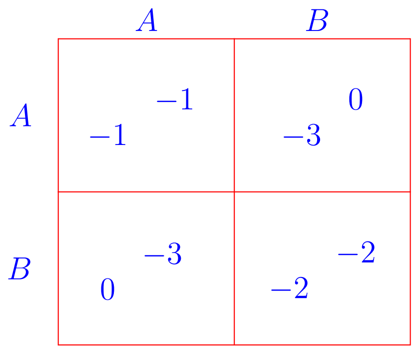
- In classical ascending price auction (ebay) it is "obvious" for a bidder how to play (accept a price if below what I want to pay).
- In a 2^{nd}-price (sealed bid) auction, that is perhaps not so obvious (to be convinced to write my true valuation I need to understand the proof that Vickrey auction is truthful).

In this lecture we consider *obviously strategyproof* mechanisms which capture this type of issues, and see some example of such mechanisms (Stable Matching and Cost-Sharing problems).

1 Dominant and Obviously Dominant Strategies

Even the Prisoners' Dilemma requires some "non-obvious" reasoning:

> "*...if the other player chooses A then I'm better choosing B because I get* 0 *instead of* -1*; If the other player chooses A...*"



Prisoners' Dilemma

In the following game, things are much simpler for the players:

	A	B
negative utility $\Leftarrow A$	-1, -1	-3, 0
positive utility $\Leftarrow B$	1, -3	0, 0

Modified Prisoners' Dilemma

"*…B gives always a nonnegative utility, while A gives a negative utility.*"

Definition 1 (dominant strategy)**.** *A strategy s_i^* is dominant if, for any s_i',*

$$u_i(s_i^*, s_{-i}) \geq u_i(s_i', s_{-i}) \qquad \text{for all } s_{-i}. \tag{1}$$

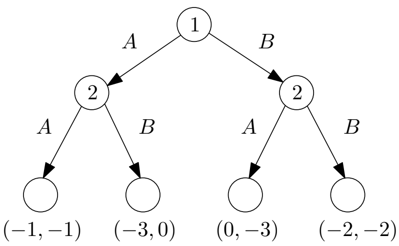
Definition 2 (obviously dominant strategy)**.** *A strategy s_i^* is obviously dominant if, for any s_i',*

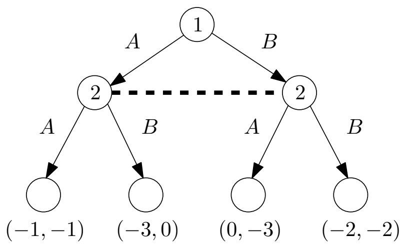
$$\min_{s_{-i}^-} u_i(s_i^*, s_{-i}^-) \geq \max_{s_{-i}^+} u_i(s_i', s_{-i}^+)$$

These definitions apply to strategic games where players move simultaneously.

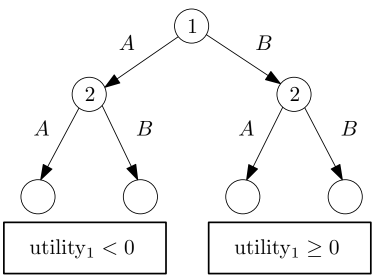
2 Extensive Form Games

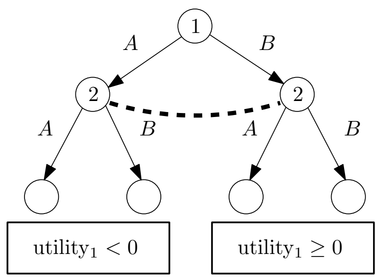
We shall consider mechanisms which involve a sequence of moves by the players (like best-response mechanisms). Consider the Prisoners' Dilemma in which first player 1 moves and then player 2 moves, forming the following trees:





The dashed line in the right tree means that player 2 does not know what player 1 has done (therefore 2 chooses A in both cases or B in both cases). Now the Modified Prisoners' Dilemma shows intuitively what obviously dominant means:





Extensive Form Games (intuitive definitions)

- History (the actions taken so far);
- Information set (what a player knows about history);
- Actions (available to the players at a given point – history);
- Strategy (choose an action based on the information)

We next translate (extend) the definition of obviously dominant strategy in the previous section (Definition 2) to extensive form games, where the strategy must take into account the history and the information available to player i. To make the presentation simpler, we assume players have complete information.

Extensive Form Games (Complete Information)

In an extensive form game H we have:

- A tree in which each node h is called a history (the root is the empty history).
- The leaves are called *terminal nodes* and correspond to the possible outcomes of the game.
- For each non-terminal history h we have
  - A player $i(h)$ that moves at this point;
  - A set of actions $A(h)$ available to this player at this point;
  - Each action $a \in A(h)$ brings to a child node $h' = succ(h, a)$.
  - In the complete information setting, player $i(h)$ knows h.
- The strategies $s = (s_1, \ldots, s_n)$ of the players (actions chosen at each h) determine the outcome $H(s_1, \ldots, s_n)$ of the game, and the utility $u_i(s_1, \ldots, s_n)$ of each player i. We sometimes consider the utilities when the game starts at some history h, and write $u_i(s_1, \ldots, s_n|h)$.

The strategy of a player i is a function s_i which maps any h in which i moves into one of the available actions at this point.

Complete Information Setting: $s_i(h) \in A(h)$

Note that each player i could move several times during the game. The set of all possible histories in which i moves and the resulting available actions that can be taken during the game are:

$$H_i = \{h | i(h) = i\}, \qquad \mathcal{A}_i := \{A(h) |\ h \in H_i\} \ .$$

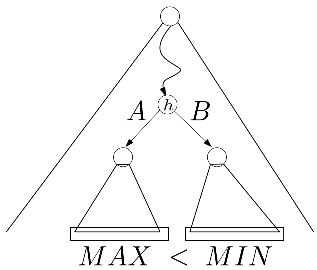
Definition 3. *A strategy for player i in an extensive form game with complete information is a function $s_i : H_i \to \mathcal{A}_i$ such that $s_i(h) \in A(h)$.*

3 Obviously Strategyproof Mechanisms

We say that two strategies s_i and s'_i *diverge at h* if

$$s_i(h) \neq s'_i(h)$$

and node h can be reached in two profiles (s_i, s_{-i}) and (s'_i, s_{-i}), for some s_{-i}.



Definition 4 (obviously dominant). *A strategy s^*_i is obviously dominant if, for any s'_i the following holds. For any h such that s_i and s'_i diverge at h, it holds that*

$$\min_{s^-_{-i}} u_i(s^*_i, s^-_{-i}|h) \geq \max_{s^+_{-i}} u_i(s'_i, s^+_{-i}|h) \ .$$

Intuitively, we say that a mechanism is obviously strategyproof if it can be implemented as an extensive form game with obviously dominant strategies. We think of a setting in which each player has a private valuation v_i or a private rank $\prec_i$ over the outcomes. In either case, we simply talk about the **private type** θ_i of i.

Definition 5 (obviously strategyproof mechanism). *A mechanism M is obviously strategyproof if there exists an extensive form game H with strategies $s_i^{\theta_i}$ such that, for all i and all θ_i,*

1. $M(\theta_1, \ldots, \theta_i \ldots, \theta_n) = H(s_1^{\theta_1}, \ldots, s_i^{\theta_i}, \ldots, s_n^{\theta_n})$;
2. $s_i^{\theta_i}$ *is obviously dominant in H.*

Remark 1. *Note that strategyproof and truthful are synonymous. Truthful means that truth-telling is a dominant strategy.*

4 Interns-Hospitals Matching

Recall that in this restriction of stable matching, the hospitals have a **common rank** $\succ_{hosp}$ over the interns. By renaming the interns, we can assume this order be just

$$1 \succ_{hosp} 2 \succ_{hosp} \cdots \succ_{hosp} n \ . \quad (2)$$

The interns (players) have their private preferences over the hospitals. The best-response mechanism (interns proposal) which makes players play in this particular order is obviously strategyproof.

Interns-Proposal Mechanism:

1. At step i, intern i checks which hospitals are not yet taken, and he/she proposes to his/her most preferred one in this set.

2. Each hospital accepts the most preferred intern that proposing to it (and this hospital is then considered taken).

Theorem 6. *The Iterns-Proposal Mechanism is obviously strategyproof.*

Proof. When player (intern) i has to make a choice, we are in a history h with available actions (non-taken hospitals)

$$A(h) = \{j_1, j_2, \ldots, j_k\}$$

Given i's preference ($\prec_i$) we let $u_i^{max}(h)$ be the utility of i if matched with the most preferred hospital j^* in $A(h)$. Observe that

1. The strategy s_i^* which makes i proposing to such j^* guarantees i being matched to j^* no matter the strategies of the other players.

2. Any strategy s_i' which deviates from s_i^* at some h will only give i a worse utility (either being matched to a worse hospital in $A(h)$ or being unmatched if proposing to a previously taken hospital).

That is, for any s_{-i}^-

$$u_i(s_i^*, s_{-i}^-|h) = u_i^{\max}(h)$$

while for any s_{-i}^+

$$u_i(s_i', s_{-i}^+|h) < u_i^{\max}(h) \quad .$$

This shows that the strategy s_i^* in Step 1 of the Interns-Proposal Mechanism is obviously dominant (see Definition 4). □

5 Single Item Auctions

Consider an **ascending price** auction for selling a single item to a set of n bidders as follows:

Ascending Price Auction:

1. Start with the set S of all bidders, and initial prices $p_i^0 = 0$.

2. At time step t, increase the price of a bidder i in the current set S.

3. Drop bidder i from the current set if $p_i^t > b_i$.

Theorem 7. *Ascending Price Auctions are obviously strategyproof.*

Proof. The last step of the mechanism corresponds to the strategy $s_i^{v_i}$ of accepting every price not larger than v_i, and rejecting every price above v_i. Since we never accept prices above v_i, we are guaranteed a non-negative utility. That is, for any s_{-i}^-,

$$u_i(s_i^{v_i}, s_{-i}^-|h) \geq 0 \ .$$

At any point h in which s_i' deviates from $s_i^{v_i}$ either (1) we accept a price higher than v_i or (2) we reject a price at most v_i. In either case, for any s_{-i}^+,

$$u_i(s_i^{v_i}, s_{-i}^+|h) \leq 0$$

were for case (1) we use that prices never decrease over time (at later steps s_i' may reject some price or keep accepting prices which are not lower than the price accepted at h). □

6 Cost-Sharing

Consider the following game (see Figure 1). The players are located on the lower node, and

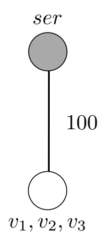


Figure 1: A simple cost-sharing problem.

they would like to be connected to the root (server) in order to receive some information (e.g., streaming of a movie). The server incurs a cost represented by the weight of the link between the two nodes if one or more players get connected (otherwise there is no cost). The **social welfare** of a solution S (which players get connected) is

$$SW(S, v) = \sum_{i \in S} v_i - C(S)$$

where $C(S) = 100$ if S is non-empty, and $C(\emptyset) = 0$.

Example 8 (VCG mechanism)**.** *The following VCG mechanism maximizes the social welfare and satisfies voluntary partecipation:*

1. *Compute S^* maximizing $SW(S, b)$, where $b = (b_1, \ldots, b_n)$ are the bids of the players.*
2. *Charge each $i \in S^*$ an amount*

$$P_i^{VCG}(b) = SW(S_{-i}^*, b_{-i}) - (SW(S^*, b) - b_i)$$

where S_{-i}^* maximizes the social welfare for the instance in which i is not present.

Exercise 1. *Prove that for the game in Figure 1 the VCG mechanism may provide the service for free, that is, for some $v = (v_1, v_2, v_3)$, each user gets connected and pays nothing.*

We are interesting in mechanisms satisfying the following condition:

Budget-balance
Sum of Payments = Cost

We consider the following class of so-called **cost-sharing** problems in which we have:

- A set of players N interested in some service;
- A cost function $C(\cdot)$ which specifies the cost $C(S)$ of providing the service to S.

Each player i has a private valuation v_i for the service and can make a bid b_i to the mechanism; The utility of players i getting the service is v_i minus their payment to the mechanism.

The main idea of the next mechanism is to divide the cost $C(S)$ among the players in S using a so called **cost-sharing method** σ which specifies the price $\sigma_i(S)$ that player $i \in S$ should pay if S are the serviced players.

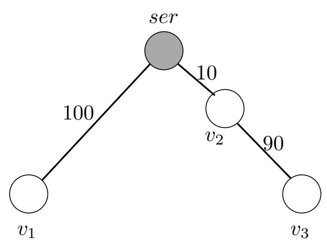
Cost-sharing Mechanism M_σ

1. Start with the set S of all players;
2. If there is a player i in S with $b_i < \sigma_i(S)$ then drop i from S; Repeat this step, in any order, until all i in S satisfy $b_i \geq \sigma_i(S)$.
3. Service all players in S and charge each $i \in S$ an amount $\sigma_i(S)$.

Exercise 2. *Consider the mechanism M_σ where σ* ***divides the cost equally*** *among all serviced players:*

$$\sigma_i(S) = \frac{C(S)}{|S|}$$

Show that the mechanism is ***not truthful*** *in the following cost-sharing game:*



where the cost $C(S)$ is the cheapest tree connecting all $i \in S$ to the root (server).

It is crucial for the mechanism to work that prices never decrease over time. In particular, one needs σ to satisfy the following condition:

Definition 9. *A cost-sharing method σ is cross monotonic if, for any S, S' with $S \subset S'$ it holds*

$$\sigma_i(S) \geq \sigma_i(S') \qquad \text{for all } i \in S$$

Theorem 10. *For any cross monotonic cost-sharing method σ, mechanism M_σ is obviously strategyproof.*

Proof. The mechanism can be seen as an ascending price auction (the item is the service and each player either wins the item or not). Indeed, at iteration t we have a subset S^t of players with

$$S^0 \supset S^1 \supset S^2 \supset \cdots \supset S^t \supset \cdots$$

and by cross- monotonicity for any i present until step t we have

$$\sigma_i(S^0) \leq \sigma_i(S^1) \leq \cdots \leq \sigma_i(S^t)$$

□

Exercise 3. *Give a direct proof that M_σ is strategyproof (truthful) for any cross-monotonic cost-sharing method σ.*

6.1 Application: Cost-Sharing on Fixed Trees

We are given a rooted weighted tree with each node i containing a player i willing to pay v_i for being connected to the root. The cost $C(S)$ for connecting a subset S of players is the cost of the cheapest rooted subtree containing all nodes S.

We use the following *fair cost-sharing* scheme, in which the cost C_e of every edge e is shared equally among all players using that edge to reach the root. For any player i, there is a unique path $\mathcal{P}_i$ from its node to the root. For any S, let $n_e(S)$ be the number of players whose path contains e. Then the cost-sharing method is

$$\sigma_i(S) = \sum_{e \in \mathcal{P}_i} \frac{C_e}{n_e(S)} \tag{3}$$

Observation 11. *The cost-sharing method in (3) is cross-monotonic, and the sum of all $\sigma_i(S)$ for $i \in S$ is exactly $C(S)$.*

The resulting mechanism is obviously strategyproof and it also satisfies the following natural conditions:

Voluntary participation: Truth-telling players never pay more than their valuation;

Consumer sovereignty: Each player can get the service if bidding high enough;

No positive transfer: Players do not pay the mechanism.

Budget Balance: The sum of the payments is equal to the cost of the service players.

Final Remarks and Recommended Literature

We have assumed players have complete information. Formally this would mean the following for our two applications:

- *Matching*: Each intern not only knows the still available hospitals, but knows precisely all previous proposals made by the other interns, and which of these have been accepted by the hospitals.
- *Auctions:* Each bidder knows exactly all prices offered to the others in the previous steps, and who accepted and who rejected.

Note that this additional information is irrelevant for the final utility of the players and therefore this assumption can be easily dropped by observing that two histories with the same information lead to the same utilities if the same action is taken.

The notion of *obviously dominant* and *obviously strategyproof* mechanism has been formally introduced here for the more general class of extensive form games with *partial information*:

- Shengwu Li. Obviously Strategy-Proof Mechanisms. Available at http://ssrn.com/abstract=2560028

The Interns-Hospital obviously strategyproof mechanism (with other impossibility results) is described in

- Itai Ashlagi and Yannai A. Gonczarowski. No stable matching mechanism is obviously strategy-proof. arXiv preprint arXiv:1511.00452 (2015).

Deferred acceptance mechanisms are a general technique to design obviously strategyproof mechanisms for combinatorial auctions:

- Paul Milgrom and Ilya Segal. Deferred-acceptance auctions and radio spectrum reallocation. EC, 2014.
- Paul Dütting, Vasilis Gkatzelis, and Tim Roughgarden. The Performance of Deferred-Acceptance Auctions. EC 2014.

These mechanisms use an idea similar to the cost-sharing mechanism M_σ discussed in this lecture. For cost-sharing problems see Chapter 15 of the AGT book.