Truthful Single-Parameter Mechanisms

ETH Zürich — Paul Dütting, Peter Widmayer

We have seen that the Vickrey (second-price) auction has several desirable properties: it incentivizes bidders to bid truthfully, it assigns the item to the bidder with the highest value, it can be computed in polynomial time.

We will define a general class of mechanism design problems; these will be characterized by the fact that each agent's private information (its value in the case of the second-price auction) consists of a single number. For this class of problems we will present a very handy characterization—known as Myerson's Lemma—of the things that we can/cannot achieve in a truthful mechanism.

1 Single-Parameter Mechanisms

In a *single-parameter mechanism design problem* a set $\mathcal{N}$ of n players (or agents) interacts with a mechanism to select a feasible outcome. Each agent has a private type $\theta_i \in \Theta_i$, which describes her preferences over outcomes. Feasible outcomes correspond to n-dimensional vectors $x \in X$, where $x_i \in \mathbb{R}$ denotes the part of the outcome that player i is interested in. We will restrict attention to settings where the type of agent i can be interpreted as a value v_i or a cost c_i; and player i's value or cost for outcome x is $v_i \cdot x_i$ or $c_i \cdot x_i$.

A (direct) mechanism $M = (f, p)$ consists of an outcome rule $f : \Theta \to X$ and a payment rule $p : \Theta \to X$.[1] A mechanism asks the players to report their types, which we will denote by b. Think of b_i as player i's bid, reported value in settings where we talk about values and reported cost when players have costs. In the former, $p_i(b)$ will be the payment that the player has to make to the mechanism, in the latter $p_i(b)$ is the payment that the mechanism makes to the player. We make the standard assumption of quasi-linear utilities. That is, player i's utility in the value-case is $u_i^M(b, \theta_i) = v_i \cdot f_i(b) - p_i(b)$ and it is $u_i^M(b, \theta_i) = p_i(b) - c_i \cdot f_i(b)$ when we talk about costs. When it is clear from the context, which mechanism we are referring to we will drop the superscript M.

The basic dilemma of mechanism design is that the mechanism designer (think of a government or company) wants to optimize some global objective such as the social welfare $\sum_{i \in \mathcal{N}} v_i \cdot x_i(b)$ by computing an allocation x based on the bids b, while the players choose their bids b_i so as to maximize their utilities $u_i(b, \theta_i)$.

Example 6.1 (Single-Item Auction)**.** *In a single-item auction n bidders compete for the assignment of an item. Each player can get the item or not, so $X_i = \{0, 1\}$, where we interpret $x_i = 1$ as bidder i gets the item. Then feasible assignments are vectors $x \in X \subseteq \prod_i X_i = \{0, 1\}^n$ with $\sum_i x_i = 1$. Each bidder i has a private value v_i for the item. Our goal is to allocate the item to the bidder with the highest value.*

Example 6.2 (Sponsored Search Auction)**.** *In a sponsored search auction we have n bidders and k positions. Each position has an associated click-through rate α_j, where we assume that positions are sorted such that $\alpha_1 > \alpha_2 > \dots > \alpha_k > 0$. Feasible allocations are $x \in X$ for which $x_i \in X_i = \{0, \alpha_k, \dots, \alpha_1\}$ for all i and for $i \neq j$ we can only have $x_i = x_j > 0$ if $x_i = x_j = 0$. Our goal is to maximize social welfare.*

[1] Such mechanisms are called direct because they simply ask the players' for their types; we will talk about indirect mechanisms later on.

Example 6.3 (Scheduling on Related Machines). *There are n machines, and each player has a private speed s_i. The inverse of the speed $t_i = 1/s_i$ is the time that machine i takes to process a job of unit length. There are m jobs with loads $\ell_1, \dots, \ell_m$, which need to be allocated to the machines. An allocation induces a work load $W_1, \dots, W_n$ for each machine. Each machine is interested in maximizing $u_i(b) = p_i(b) - W_i \cdot t_i$, while the mechanism designer wants to minimize the makespan $\max_i W_i \cdot t_i$.*

A very elegant way to resolve the potential conflict of interest between the mechanism designer and the players, is to ensure that it is in each player's interested to bid truthfully. In this case $b_i = \theta_i$ for all $i \in N$ and by choosing an outcome that is optimal for b the mechanism designer chooses an outcome that is optimal for v.

Definition 6.4. *A mechanism $M = (f, p)$ is called dominant strategy incentive compatible (DSIC) (or just truthful), if for each player i bidding $b_i = \theta_i$ is a weakly dominant strategy. That is, for all $i \in \mathcal{N}$, $\theta_i \in \Theta_i$, and all $b \in \Theta$ it holds that*

$$u_i^M((\theta_i, b_{-i}), \theta_i) \geq u_i^M((b_i, b_{-i}), \theta_i) \ .$$

Our main question for now will be to identify those outcome rules f, for which we can find payment rules p such that $M = (f, p)$ is a truthful mechanism. We will call these outcome rules *implementable.*

2 Myerson's Lemma

It turns out that there is a very satisfying answer to this question, if we confine ourselves to single-parameter environments.

Definition 6.5. *An allocation rule f for a single-parameter mechanism design problem is monotone if for each player $i \in \mathcal{N}$ and for all bids b_{-i} of the players other than i, the allocation $f_i(z, b_{-i})$ to player i is non-decreasing in bid z.*

Theorem 6.6 (Myerson 1981). *For single parameter environments, the following three claims hold: (1) An allocation rule is implementable if and only if it is monotone. (2) If allocation rule f is monotone, then there exists a unique payment rule p such that the mechanism $M = (f, p)$ is truthful, assuming that a zero bid implies a zero payment. (3) The payment rule according to (2) can be expressed by an explicit formula.*

This result is remarkable for several reasons: (1) It reduces the rather abstract problem of deciding whether a certain allocation rule can be implemented, to the far more operational question of whether a given allocation rule is monotone. (2) It leaves essentially no ambiguity in regard to the payments. If we require that an agent with value zero pays nothing, then there is a unique payment rule that turns a given allocation rule into a truthful mechanism. (3) It gives an explicit formula for the payments that achieve this.

Proof. Let us consider any allocation rule f, whether monotone or not, and let us study how truthful payments could look like. Truthfulness requires that the utility of each bidder is maximized by bidding truthfully, no matter who bids and no matter what the other players' bids are, where the utility of player i for bid z is $u_i(z, b_{-i}) = v_i \cdot f_i(z, b_{-i}) - p_i(z, b_{-i})$ for b_{-i} denoting the bids of the other players.

We will now inspect what happens when player i does not bid truthfully, by comparing the resulting utility with the utility that comes from a truthful bid. Since we fix player i and we fix the bids of all other players, instead of $u_i(z, b_{-i}) = v_i \cdot f_i(z, b_{-i}) - p_i(z, b_{-i})$ we simply write $v \cdot f(z) - p(z)$. A bid that is not truthful can be either lower or higher than the true value. Let us therefore fix two possible bids y and z for player i, with $0 \leq y < z$. If z is our player's true value, her true bid results in a utility of $z \cdot f(z) - p(z)$, while the untruthful lower bid y results in a utility of $z \cdot f(y) - p(y)$. Truthfulness now requires that $z \cdot f(z) - p(z) \geq z \cdot f(y) - p(y)$. Symmetrically, if y is the true value and our player bids z, truthfulness requires $y \cdot f(y) - p(y) \geq y \cdot f(z) - p(z)$. Rearranging terms and writing both inequalities together, we get lower and upper bounds on the payment difference for both bids: $z \cdot (f(y) - f(z)) \leq p(y) - p(z) \leq y \cdot (f(y) - f(z))$, often called payment difference sandwich. The payment difference sandwich allows us to infer the forward direction of part (1) of the theorem: Since $y < z$, we must have $f(y) - f(z) \leq 0$, the monotonicity of the allocation rule.

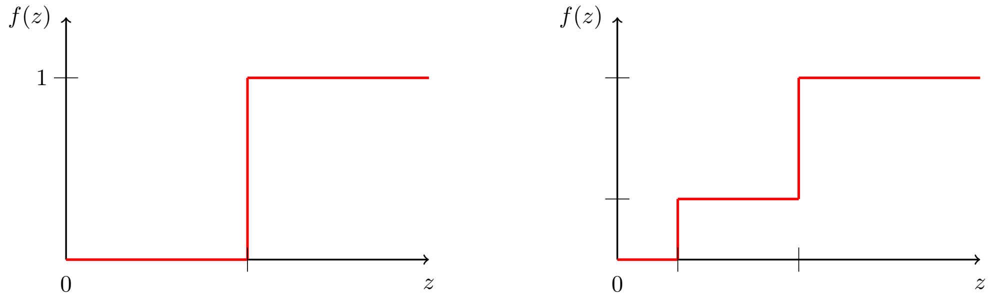

Figure 1: Piece-wise constant allocation curves

For the sake of simplicity, let us limit ourselves to allocation rules that are piecewise constant, as in the Vickrey auction of a single item, or in sponsored search. See Figure 1 for an illustration. Now let us study what happens if y and z approach each other arbitrarily. If they approach each other on a constant part of the allocation curve where there is no jump (discontinuity), both allocations are the same, and the payment difference sandwich requires that the payments be the same (since the payment difference is bounded by 0 from below and from above). If, however, both bids approach each other at the exact location of a jump, there will be an allocation difference, as defined by the height of the jump. Calling the height of the jump h, the payment difference sandwich becomes $z \cdot h \leq p(y) - p(z) \leq z \cdot h$. Again, lower and upper bounds coincide, and the payment difference (at z) is required to be z times the allocation difference (at z). Notably, any other payment difference will not allow for a truthful mechanism. For the payment function as a whole, these fully prescribed payment differences need to be summed up from zero (recall the assumption that $p(0) = 0$. For a bid b_i of bidder i and an allocation that jumps in points $z_1, \ldots, z_l$ in the closed interval from 0 and b_i, we get the payment function

$$p_i(b_i, b_{-i}) = \sum_{j=1}^{l} z_j \cdot \text{jump in } f_i(\cdot, b_{-i}) \text{ at } z_j.$$

This proves part (2) of the theorem, since it leaves no other option for the payment function to be truthful. Also, it proves part (3) at the same time. Now, the only missing step in the proof of the theorem is the truthfulness of this payment scheme.

We convince ourselves pictorially that this payment scheme is truthful, see Figure 2. In all three parts of Figure 2, the allocation curve is the same, as well as the true value of our player. Figure 2 (a) shows what happens in a truthful bid: Our bidder gets the surplus indicated by the area of the blue rectangle, with the red area showing her payment and the green area her utility. Figure 2 (b) shows what happens when she overbids: For bid b with $v < b$, her allocation goes up and therefore her surplus goes up (blue), but her pay (red) goes up by more than her surplus, resulting in a utility that is lower (the lower green L-shape minus the small green rectangle). On the other hand, underbidding (Figure 2 (c)) leads to a smaller allocation, smaller surplus (blue), smaller pay (red), but also smaller utility (green). That is, the player's utility is indeed maximized by her true bid, which proves the theorem.

□

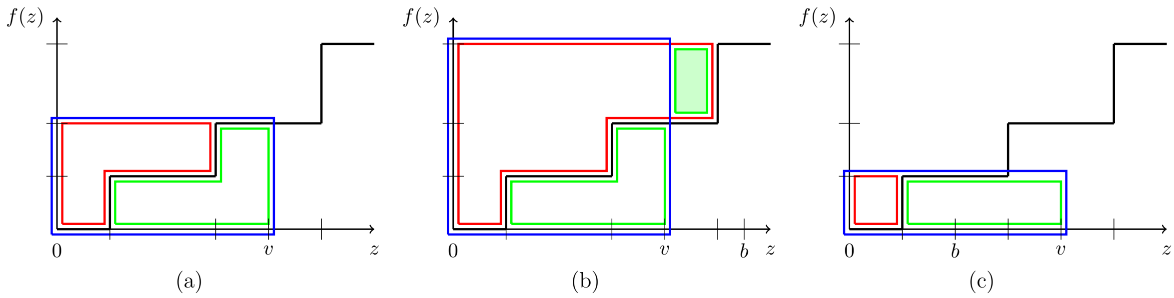

Figure 2: Visualization of the value (blue), the payment (red), and the utility (green) when bidding truthfully (on the left) and for over- and underbidding (in the middle and on the right). Shaded areas contribute negatively.

3 The VCG Mechanism

Part of the strength of Myerson's lemma stems from the fact that it applies to any objective. If we specialize the objective, we may obtain simpler answers.

Theorem 6.7 (Vickrey, Clarke, Groves 1961-73)**.** *Let f be an allocation rule that maximizes social welfare. Let x^* be an allocation that maximizes $\sum_{i \in \mathcal{N}} x_i^* \cdot b_i$, then charging each player $i \in \mathcal{N}$,*

$$p_i(b_i, b_{-i}) = \max_{x \in X} \sum_{j \neq i} x_i \cdot b_j - \sum_{j \neq i} x_j^* \cdot b_j$$

is a truthful mechanism.

The payments can be thought of player i's "externality", the amount by which his presence reduces the combined welfare of the other players.

Proof. We first observe that the allocation rule f that maximizes social welfare is monotone. To see this assume by contradiction that for some player i and a pair of bids $b_i < b'_i$ we have $f_i(b_i, b_{-i}) > f_i(b'_i, b_{-i})$. Denote the allocation that results at $b = (b_i, b_{-i})$ by x and the allocation that results at $b' = (b'_i, b_{-i})$ by x'. By optimality of x and x',

$$x_i b_i + \sum_{j \neq i} x_j b_j \geq x'_i b_i + \sum_{j \neq i} x'_j b_j \ ,$$
$$x'_i b'_i + \sum_{j \neq i} x'_j b_j \geq x_i b'_i + \sum_{j \neq i} x_j b_j \ .$$

By adding the first inequality to the second and cancelling the sums, we obtain

$$x_i b_i + x'_i b'_i \geq x'_i b_i + x'_i b_i \ \Leftrightarrow \ x_i(b_i - b'_i) \geq x'_i(b_i - b'_i) \ \Leftrightarrow \ x_i(b'_i - b_i) \leq x'_i(b'_i - b_i) \ ,$$

where we had to reverse the inequality in the final step because $b_i - b'_i < 0$. Since $b'_i > b_i$ we can divide through $b'_i - b_i > 0$ and conclude that $x_i \leq x'_i$, but this contradicts our assumption that $x_i = f_i(b_i, b_{-i}) > f_i(b'_i, b_{-i}) = x'_i$.

Having shown that the allocation rule is monotone we can apply Myerson's Lemma to get a formula for the payments, which we need to show is identical to the formula for the payments given in the statement of the theorem. In the remainder we will again assume that the allocation curves $f_i(\cdot, b_{-i})$ are step functions.

We will in fact restrict attention to the case where the allocation curve of player i has a single jump from 0 to 1 at some point z. A similar argument applies to the case where there can be multiple jumps. By Myerson's lemma the payment $p_i(b_i, b_{-i})$ of player i is zero if $b_i < z$ and it is z otherwise. Suppose $b_i \geq z$. At z the welfare maximizing allocation changes from

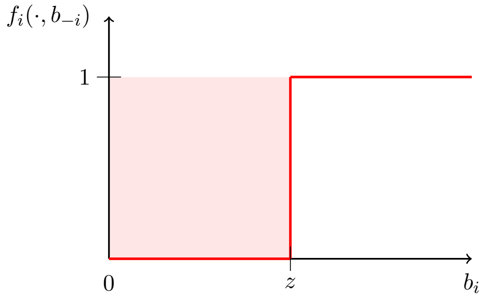

Figure 3: Allocation curve with a single step from 0 to 1 at $b_i = z$

$\max_{x \in X} \sum_{j \neq i} x_j b_j$ (for points to the left of z) to $x_i^* y + \sum_{j \neq i} x_j^* b_j$ (for points y to the right of z) so that at $y = z$ we have

$$\max_{x \in X} \sum_{j \neq i} x_j b_j = x_i^* z + \sum_{j \neq i} x_j^* b_j.$$

Noting that $x^* = 1$ and subtracting $\sum_{j \neq i} x_j^* b_j$ from both sides we get the desired formula for $p_i(b_i, b_{-i}) = z$. □

Alternatively, one could prove that f together with p as described in the theorem is a truthful mechanism, by arguing directly that in the induced mechanism $M = (f, p)$ bidding truthfully is a weakly dominant strategy for each player.

4 Examples

We are now ready to apply the tools that we developed in this lecture to the three examples mentioned at the beginning.

Example 6.8 (Single-Item Auction)**.** *We have already seen that the Vickrey (second-price) auction is truthful. We can recover this result from both Myerson's lemma and the VCG result. From Myerson's lemma we know that the payment for winning is the critical value at which a player becomes a winner. This is the second highest bid. From the VCG result we know that the winner's payment is equal to his "externality". If the winner is not present the bidder with the second highest bid will win, if he is present then he wins and the others get nothing. So his payment is equal to the second-highest bid.*

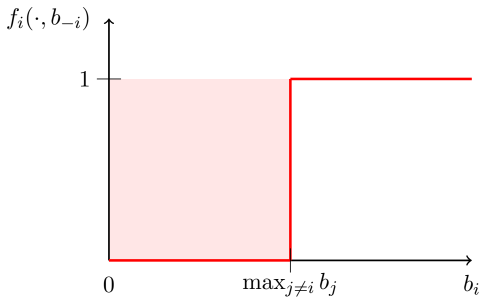

Figure 4: Allocation curve in a single-item auction

Example 6.9 (Sponsored Search Auction)**.** *In sponsored search social welfare is maximized by greedily assigning position* 1 *through* k *to the bidders with the* 1*-st to* k*-th highest bid. Denoting the* j*-th highest bid by* $b_{(j)}$*, Myerson's lemma yields the following graphical representation of a player's payment whose bid is highest:*

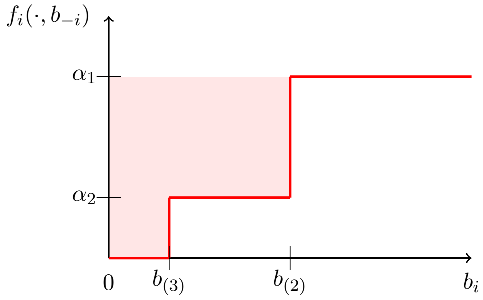

Figure 5: Allocation curve in a sponsored search auction

More generally, the externality of a player i that is assigned position j is the loss in welfare incurred on the players assigned slots below. If player i was not present they could all move one position up. In other words, setting $\alpha_{k+1} = 0$, player i's payment is given by

$$p_i(b_i, b_{-i}) = \sum_{\ell=j}^{k} (\alpha_j - \alpha_{j+1}) \cdot b_{(j+1)} \quad .$$

Example 6.10 (Scheduling on Related Machines)**.** *Efficiently minimizing the makespan on related machines is impossible (unless $P = NP$). So let's consider the following simple algorithmic strategy: Arrange the jobs in decreasing work order, and then greedily assign the next job to the machine that finishes it earliest. For example, three jobs with* 2, 1.1, *and* 1.05 *work units will be placed on two machines with* 0.4 *and* 0.45 *processing times per unit of work as follows. The heaviest job will go to the faster machine, and the other two lighter jobs will go to the slower machine. Now consider what happens if the slower machine claims to have a processing time of* 0.5*. Then it will receive the heaviest job, while the two smaller jobs are assigned to the other machine. So by claiming a higher speed (shorter processing time), a machine can reduce its workload. A contradiction to monotonicity.*

Let us conclude with two important orthogonal observations: (1) In many practical applications to which Myerson's Lemma applies, other (non-truthful) mechanisms are used in practice. For example, the mechanism used by Google to sell sponsored search results is not the VCG mechanism. So there must be other reasons, in addition to truthfulness, that play a role. We will return to this point and non-truthful mechanisms later. (2) Myerson's lemma tells us that we can find the best truthful polynomial-time mechanism for a problem by searching for the best polynomial-time algorithm that is *monotone*. An important question thus is, does this additional requirement make the problem any harder? We will discuss this tradeoff between incentives and computation in the next lecture.

Recommended Literature

- Tim Roughgarden's lecture notes `http://theory.stanford.edu/~tim/f13/l/l3.pdf`
- R. Myerson, Optimal Mechnism Design, Mathematics of Operations Research, 6:58–73, 1981. (Original characterization of truthful mechanisms)
- A. Archer and É. Tardos, Truthful Mechanisms for One-Parameter Agents. FOCS 2001. (Characterization of truthful mechanisms, which is deemed more accessible to computer scientists)