

## 20.1 Introduction

Today we're finally going to go beyond single-parameter environments. We will define and analyze the main result from classical mechanism design, known as the VCG mechanism.

## 20.2 Setup & Examples

### 20.2.1 General Setup

- Bidders/agents/players  $[n]$
- Finite set  $\Omega$  of *outcomes* (think allocations)
- Each player  $i$  has a valuation function  $v_i : \Omega \rightarrow \mathbb{R}_{\geq 0}$
- The social welfare of some outcome  $w \in \Omega$  is  $\sum_{i=1}^n v_i(w)$ .

### 20.2.2 Examples

**Single Item:**  $\Omega$  is the  $n + 1$  possible allocations ( $n$  different winners or no winner), and  $v_i(w)$  is equal to  $v_i$  if  $i$  wins the item in  $w$  and 0 otherwise. Note that in the more general setting, we can have more interesting valuation functions even for a single item. Maybe player  $i$  would like to get the item, but if they don't then they would still prefer player  $j$  to get the item over player  $k$ . We can model that now by an appropriate valuation function  $v_i$ !

**Single-Parameter Environments:**  $\Omega = X$  (the outcomes are the feasible allocations), and  $v_i(w) = v_i w_i$  (here  $w$  is an  $n$ -dimensional vector since  $\Omega = X$ ).

**Combinatorial Auctions:** We'll talk a lot more about these next lecture. But they're intuitive and good to keep in mind for general auctions, so I'll give the definition here.

- Players  $[n]$ .
- $m$  items  $M$ .
- Each player  $i$  has a valuation for each "bundle"  $S \subseteq M$ , so  $v_i : 2^M \rightarrow \mathbb{R}$  (note that there are  $2^m$  private parameters now instead of just one!)
- The outcomes are allocations  $f : M \rightarrow [n]$ , so  $v_i(f) = v_i(\{j : f(j) = i\})$ .

## 20.3 Main Result

Today we're going to try to prove the following result:

**Theorem 20.3.1** (Vickrey-Clarke-Groves). *In every general mechanism design environment, there is an incentive-compatible mechanism which maximizes social welfare.*

Clearly this is a very powerful theorem! But there are a few subtleties and caveats that we should mention before we actually prove it. First, this is not an “awesome” auction: we cannot compute it in polynomial time. This should not be surprising, since we cannot efficiently compute the welfare-maximizing allocation even in single-parameter environments like knapsack auctions.

Second, for the first time in a couple weeks we are not able to use Myerson to get prices once we've decided how to choose an outcome (since we're no longer in a single-parameter environment). So we'll have to argue incentive compatibility directly.

Third, we have to be a bit careful about what we mean by a “bid” in this context. The obvious thing, which is what we're going to do, is to let a “bid” consist of all the private information, i.e., bidder  $i$  sends a bid *function*  $b_i : \Omega \rightarrow \mathbb{R}_{\geq 0}$ . Then a mechanism is incentive compatible if truthfully sending  $b_i = v_i$  is a dominant strategy for player  $i$ . This clearly has some issues in practice, since in general  $v_i$  can be huge, and there's a lot of interesting work on preference elicitation, bidding languages, and other extensions of the setting to enable bids to be a little more reasonable. But that's what we'll use for today.

### 20.3.1 The VCG Mechanism

To define the mechanism we need to define the allocation/outcome rule and the prices to charge. Like always, we'll assume that the utility of a player is its valuation minus the price it pays. Like in the single-parameter environment, we'll try to use a two-stage design process: first figure out the outcome rule, and then figure out a pricing rule to make this incentive compatible.

**Outcome rule:** We don't really have much choice here. We need to maximize social welfare, so we're basically forced to use the outcome rule which does this. Formally, given bids  $b$  (where  $b_i : \Omega \rightarrow \mathbb{R}_{\geq 0}$ ), we'll choose the outcome

$$\mathbf{x}(b) = \arg \max_{w \in \Omega} \sum_{i=1}^n b_i(w)$$

**Pricing rule:** Now comes the tricky part: how should we design prices to make this incentive compatible (and is that even possible)?

- In the single-parameter setting, we proved (Myerson's Lemma) that such a pricing function exists if the allocation rule is monotone, and then proved that the welfare-maximizing allocation rule is monotone. But what does monotone even mean here? What if there's one parameter where we over bid and one parameter where we underbid – should we get more or less value from this?

- In the single-item setting, we showed that the right price was the “critical bid”. But now even for a single-item, players are allowed to have pretty complicated valuations. What does the “critical bid” mean here?

Let’s go back to the Vickrey (second-price) auction for a single-item to get some inspiration. We’re just going to have to look at it in a slightly different way. Consider some player  $i$ . How much does player  $i$  hurt everyone else by participating in the auction?

- If  $i$  does not get the item, then it doesn’t hurt anyone – the only difference between it participating and not participating is that the winner might be charged a higher price, but the social welfare is unchanged. So the social welfare of everyone else is also unchanged.
- If  $i$  does get the item, then it was the highest bid. So everyone else has social welfare 0, i.e.,  $\sum_{j \neq i} v_j x_j = 0$ . But if  $i$  did not participate, then the item would have gone to the bidder with the highest valuation other than  $i$ , i.e., the second-highest bid. So in that case, we would have that  $\sum_{j \neq i} v_j x_j = \arg \max_{j \neq i} v_j$ . So player  $i$ , by participating in the auction, made the combined social welfare of all other players go from  $\arg \max_{j \neq i} v_j$  to 0. So the “total harm” caused by player  $i$  is the second highest bid, which is the price that player  $i$  was charged!

We’re going to generalize this idea to the general mechanism design setting: we’ll charge player  $i$  the *eternality* they cause. Formally, given bids  $b$ , let  $w^*$  be the welfare-maximizing outcome. Then the price we’ll charge player  $i$  is

$$p_i(b) = \max_{w \in \Omega} \sum_{j \neq i} b_j(w) - \sum_{j \neq i} b_j(w^*).$$

Note that this is precisely the outcome maximizing social welfare without player  $i$  minus the social welfare of the other players when player  $i$  participates, so it is the harm that player  $i$  causes the other players.

This pricing rule, combined with the social welfare-maximizing outcome rule, is the VCG Mechanism.

### 20.3.2 Analysis

Clearly VCG maximizes social welfare (by definition), so we just need to prove that it is incentive compatible, i.e., bidding  $b_i = v_i$  is a dominant strategy for all players  $i$ . So fix some player  $i \in [n]$  and other bids  $b_{-i}$ .

The utility of player  $i$  under bids  $b$  is  $v_i(w^*) - p_i(b)$ , where  $w^* = \arg \max_{w \in \Omega} \sum_{j=1}^n b_j(w)$ . Plugging in the payment rule from the mechanism, we get that the utility of player  $i$  is

$$\begin{aligned} v_i(w^*) - p_i(b) &= v_i(w^*) - \max_{w \in \Omega} \sum_{j \neq i} b_j(w) + \sum_{j \neq i} b_j(w^*) \\ &= \left( v_i(w^*) + \sum_{j \neq i} b_j(w^*) \right) - \max_{w \in \Omega} \sum_{j \neq i} b_j(w) \end{aligned}$$

The last term here is independent of player  $i$ 's bid  $b_i$ : nothing  $i$  does can affect this term. So player  $i$  is trying to maximize  $v_i(w^*) + \sum_{j \neq i} b_j(w^*)$ , which it can affect since  $b_i$  will affect  $w^*$ . What is this sum? If  $b_i = v_i$  then it is  $\sum_{j \in [n]} b_j(w^*)$ , which is exactly what the outcome rule is trying to maximize! So if player  $i$  bids  $v_i$ , then the outcome rule is choosing the outcome that is best for player  $i$ . If player  $i$  bids some  $b_i \neq v_i$ , then the outcome rule chooses some  $w^*$  that maximizes social welfare for  $b_i$  rather than for  $v_i$ , which is clearly worse for player  $i$ . Thus bidding  $b_i$  is a dominant strategy.

To see this intuitively, note that what the VCG mechanism is doing is “aligning incentives”: because of the price chosen, in order to maximize their own utility player  $i$  wants to maximize total social welfare. Since that's what the mechanism is trying to do, the best thing for player  $i$  to do is be truthful and let the mechanism do this.

So we're almost done. We just need to prove that utilities are never negative for a truthful bidder and that prices are never negative. Clearly  $p_i(b)$  is never negative, since the  $w^*$  is one particular instantiation of the max. To prove nonnegative utilities, let  $w' = \arg \max_{w \in \Omega} \sum_{j \neq i} b_j(w)$ . Then the price is

$$\begin{aligned} p_i(b) &= \sum_{j \neq i} b_j(w') - \sum_{j \neq i} b_j(w^*) \\ &= b_i(w^*) - \left( \sum_{j=1}^n b_j(w^*) - \sum_{j \neq i} b_j(w') \right). \end{aligned}$$

The first term here is what player  $i$  bid, so for a truthful bidder this will be  $v_i(w^*)$ . Thus if  $i$  is truthful, it will have utility

$$v_i(w^*) - p_i(b) = \sum_{j=1}^n b_j(w^*) - \sum_{j \neq i} b_j(w') \geq \sum_{j=1}^n b_j(w') - \sum_{j \neq i} b_j(w') = b_i(w') \geq 0$$