CS880: Algorithmic Mechanism Design

Lecture 2: Introduction to Revenue Maximization **Scribe:** Austen Fan

2.1 Borodin & Lucier's Algorithm

Recall last time we introduced the greedy algorithm by Borodin and Lucier which can be thought of as a generalization of the LOS algorithm:

Algorithm 2.1.1 (Borodin & Lucier '2011)

- *Ask agents to report valuation* $\tilde{v_i}(\cdot)$
- *While items remain unallocated*
  - *Let* $(i, X_i) = \underset{\substack{(i,S) \\ S \text{ is unallocated}}}{arg\,max}\ f_i\left(\tilde{v_i}\left(\cdot\right), S\right)$
  - *Allocate* X_i *to agent* i
- *Charge each agent its critical value* $\theta_{i,S}$*, i.e. the min value* $\tilde{v_i}(S)$ *that* i *needs to report for* S *to be allocated* S

The function f_i in the algorithm is a "score function" that assigns a non-negative value to each pair $(\tilde{v_i}, S)$ where S is a subset of items. In LOS algorithm (single-minded agent setting), the function $f_i = \frac{v_i}{\sqrt{S_i}}$. Typically under single-minded agent setting, when f_i's are monotone non-decreasing in v_i and monotone non-increasing in $|S|$ along with charging critical price, the algorithm under the LOS framework will be DSIC. Here is a much general setting where each agent could assign different values to different subset of items and the Borodin & Lucier's algorithm is no longer DSIC. Nevertheless, Borodin & Lucier's algorithm possesses good behavior when considering worst case Nash Equilibrium. Formally speaking, we are going to prove the following claim:

Claim 2.1.2 *Suppose that the function* f_i*'s are monotone non-decreading in* v_i *and the greedy algorithm is a* c*-approximation. Then in any Nash Equilibrium of the greedy mechanism, the Social Welfare is a* $(c+1)$*-approximation to the optimal Social Welfare.*

Proof: Observe that any agent will never submit a bid that is higher than her value. Indeed, if she wins with that bid, since the critical price is defined to be independent of her bid, either she will earn a utility which is equal to the case when she bid truthfully or incur a net loss. If she loses, bidding higher than her value will not bring any good. Thus we have

$$\sum_{\substack{(i,X_i) \text{ allocated} \\ \text{by greedy}}} v_i(X_i) \geq \sum_{\substack{(i,X_i) \text{ allocated} \\ \text{by greedy}}} \tilde{v_i}(X_i)$$

By assumption,

$$\sum_{\substack{(i,X_i)\text{ allocated}\\ \text{by greedy}}} \tilde{v_i}(X_i) \geq \frac{1}{c} \sum_{(i,S_i)\text{ optimal}} \tilde{v_i}(S_i)$$

Imagine changing the bid for each (i, S_i) which is allocated under optimal but not by greedy algorithm to $\theta_i(S_i) - \epsilon$ for any $\epsilon > 0$. Then the result of greedy solution will not change. Thus,

$$\begin{aligned}\sum_{(i,X_i)\in\text{ Greedy}} \tilde{v_i}(X_i) &\geq \frac{1}{c}\{ \sum_{(i,S_i)\in\text{OPT}\setminus\text{Greedy}} (\theta_i(S_i) - \epsilon) + \sum_{(i,S_i)\in\text{OPT}\cap\text{Greedy}} \theta_i(S_i)\} \\ &\geq \frac{1}{c} \sum_{(i,S_i)\in\text{OPT}} \theta_i(S_i) - O(\epsilon)\end{aligned}$$

Since the equation holds for any $\epsilon > 0$, letting ϵ go to 0

$$\sum_{(i,X_i)\in\text{ Greedy}} \tilde{v_i}(X_i) \geq \frac{1}{c} \sum_{(i,S_i)\in\text{OPT}} \theta_i(S_i) \tag{$\star$}$$

By definition, agent i's strategy is a best response to other's bids under any Nash Equilibrium. Consider alternate strategy where agent i bids $v_i(S_i)$ on the subset of items S_i. Then

$$\begin{aligned}\text{Utility from new strategy} &= v_i(S_i) - \theta_i(S_i) \\ &\leq v_i(X_i) - \theta_i(X_i)\end{aligned}$$

where X_i and S_i are subsets of items allocated to agent i under greedy algorithm and optimal solution respectively. Combining ($\star$) yields the result

$$\begin{aligned}\sum_{(i,S_i)\in\text{ OPT}} v_i(S_i) &\leq \sum_{(i,S_i)\in\text{ OPT}} v_i(X_i) + \sum_{(i,X_i)\in\text{ Greedy}} v_i(X_i) \\ &\leq c \sum_{(i,X_i)\in\text{ Greedy}} v_i(X_i) + \sum_{(i,X_i)\in\text{ Greedy}} v_i(X_i)\end{aligned}$$

which in terms of Social Welfare is equivalent to

$$\text{OPT} \leq (c+1)\text{Greedy}$$

The proof is now complete. ■

Recall the definition of Price of Anarchy last lecture, the claim is exactly saying PoA $\leq c + 1$.

2.2 Maximal in Distributional Range

We introduce an alternate approach called *Maximal in Distributional Range*. Recall that the Social Welfare Maximization on set $\mathcal{F}$ is the following problem:

- Given a feasible set $\mathcal{F}$ and value functions $v_i : \mathcal{F} \to \mathbb{R}^+$ for any $i \in [n]$ where n is the number of agents

– Find $\arg\max_{j \in \mathcal{F}} \sum_{i\in[n]} v_i(j)$

The idea is instead of considering the whole feasible set $\mathcal{F}$, we choose to consider a subset $\mathcal{F}' \subset \mathcal{F}$ where optimal over $\mathcal{F}'$ is approximate to optimal over $\mathcal{F}$. Formally speaking,

Maximum In Range mechanism

– Find $\mathcal{F}' \subset \mathcal{F}$ such that Social Welfare Maximization over $\mathcal{F}'$ is computationally easy

– Run VCG mechanism over $\mathcal{F}'$

Similarly, we might consider picking a distribution over feasible set $\mathcal{F}$ such that optimal over that distribution gives approximate result of optimal over the entire set.

Maximum In Distributional Range mechanism

– Find $\mathcal{F}' \subset \Delta^{\mathcal{F}}$ where $\Delta^{\mathcal{F}}$ is a distribution over outcomes such that Social Welfare Maximization over $\mathcal{F}'$ is computationally easy

– Run VCG mechanism over $\mathcal{F}'$

We now describe Lavi-Swamy's approach on MIDR mechanism. Let n be the number of agents, m be the number of items and $\mathcal{F}$ be all partitions of items across agents. Consider the following Relaxation Linear Programming setting which is called **Configuration LP** where $x_{i,S}$ refers to the indicator variable that set S is allocated to agent i:

$$\begin{aligned}
\max \quad & \sum_{i,S} x_{i,S} v_i(S) \\
\text{s.t.} \quad & \sum_i \sum_{j\in S} x_{j,S} \le 1 \quad \forall j \in [m] \\
& \sum_S x_{i,S} \le 1 \quad \forall i \in [n] \\
& x_{i,S} \ge 0 \quad \forall i \in [n], j \in [m]
\end{aligned}$$

Claim 2.2.1 ***Configuration LP*** *can be solved exactly in polynomial time.*

Proof: Omitted. ■

The idea is to use a separation oracle for the dual problem. Though there might be exponentially many constraints, the algorithm only needs demand queries to solve the dual LP. That is, the algorithm inductively asks agents to pick their favorite set of items at given prices and use that information to come up with new prices. The main theorem is the following:

Theorem 2.2.2 (Lavis-Swamy '05) *Suppose there is a c-approximation rounding algorithm for the* ***Configuration LP****. Then given any x feasible for the LP, we can find a distribution D over deterministic partitions such that* $SW(D) = \frac{1}{c}\ SW(x)$.

Proof: Omitted. ■

The idea is to write $\frac{1}{c}x$ as a convex combination of many deterministic partitions which are integral solutions. Denote

$$\text{ALG}(x) := \text{distribution } D$$

as in the Lavis-Swamy theorem and

$$\mathcal{F}' = \{\text{ALG}(x) : x \in \text{ feasible set of } \textbf{Configuration LP}\}$$

Then the following mechanism gives optimal over $\mathcal{F}'$:

- Solve optimal solution x^* for **Configuration LP**
- Return ALG(x^*)
- Use VCG pricing

2.3 Revenue Maximization

We now focus on maximizing the total revenue. However, things become trickier than Social Welfare. When maximizing Social Welfare, the technique usually used is to first consider allocation and payment rule, and then justify the assumption of incentive compatibility. But that will not work when maximizing the total revenue since the highest price a seller might charge depends on the value of the buyers. Besides, the optimal solution of Social Welfare is clear and this is no longer true under revenue maximization. In summary, two immediate difficulties are:

- Can't easily compare to an incentive-free optimum
- Not clear what is a good alternative benchmark

Without an *ex ante* knowledge, it is hard to even start any analysis. We thus introduce the Bayesian assumption which is widely used in economics.

Bayesian assumption: Agent i's value function is drawn from a known distribution.

The timeline for direct revelation mechanisms is: Designers and agents learn distribution from which values are drawn $\rightarrow$ Designer announces mechanism $\rightarrow$ Values are drawn $\rightarrow$ Agents report valuation $\rightarrow$ Designer announces outcome

The simplest setting is when single item auction where there is single agent with value v drawn from a distribution $\mathcal{F}$. The outcome space is determined by allocation rule $\mathbf{x}$ and payment rule $\mathbf{p}$. We consider randomized allocation rule where $\mathbf{x} \in [0, 1]$ is the probability of allocation and assume positive payment rule, i.e. $\mathbf{p} \in \mathbb{R}^+$. We make a further assumption that agent's expected utility from outcome $(\mathbf{x}, \mathbf{p})$ is simply $\mathbf{vx} - \mathbf{p}$.

The mechanism designer's problem is formalized as following:

Given distribution $\mathcal{F}$, develop functions $\mathbf{x}(\cdot), \mathbf{p}(\cdot)$ such that

- $(\mathbf{x}, \mathbf{p})$ is DSIC/BIC (defined later)
- $\mathrm{E}_{v \sim F}[p(v)]$ is maximized

The following lemma characterizes DSIC mechanism:

Lemma 2.3.1 (Myerson) *For the single item single agent problem,* $(\mathbf{x}, \mathbf{p})$ *is DSIC iff*

1. $\mathbf{x}$ *is monotone non-decreasing in* v*;*
2. *Satisfies the payment identity* $\mathbf{p}(v) = v\mathbf{x}(v) - \int_0^v \mathbf{x}(t)dt + \mathbf{p}(0)$.

Usually we can "normalize" the payment rule so that $\mathbf{p}(0) = 0$. This is just saying agent do not need to pay anything when they don't value the item at all.

2.4 General Single-parameter Linear Agents Setting

Suppose now we have n agents. Each agent has its own "type" which is given by **one** number v_i. If the outcome for each agent is $x_i \in [0, 1]$ and $p_i \in \mathbb{R}$, then the utility of linear agent i from (x_i, p_i) is given by $v_i x_i - p_i$. Each agent acts to maximize their expected utility.

The timeline for direct revelation mechanism is summarized as the following: Designer and agents learn the joint distribution $\mathcal{F}$ of values $\rightarrow$ Designer announces mechanism (x, p) $\rightarrow$ Agents realize their values $(v_1, v_2, \ldots, v_n) \sim \mathcal{F}$ $\rightarrow$ Agents report their bids $b_i = \sigma_i(v_i)$ $\rightarrow$ designer announces outcome $x(b_1, b_2, \ldots, b_n)$ and $p(b_1, b_2, \ldots, b_n)$.

We adopt the notation that $x_i(b_1, b_2, \ldots, b_n)$ and $p_i(b_1, b_2, \ldots, b_n)$ is the allocation to and payment of agent i. Analogous to Dominant Strategy Incentive Compatibility, we now define what is called Bayesian Incentive Compatibility.

Definition 2.4.1 (BIC) *A strategy* σ_i *for agent* i *maps value* v_i *to the bid* $\sigma_i(v_i)$. *A strategy tuple* $(\sigma_{1,\sigma_2,\ldots,\sigma_n})$ *is in Bayesian Nash Equilibrium if given all other agents' strategies* $\sigma_{\text{-}i}$*, for all agent* i*, values* v_i *and alternate strategies* σ_i'*,*

$$\mathop{\mathrm{E}}_{\substack{v_{\text{-}i} \sim \mathcal{F} | v_i \\ b_{\text{-}i} = \sigma_{\text{-}i}(v_{\text{-}i})}} [v_i x_i(\sigma_i(v_i), b_{\text{-}i}) - p_i(\sigma_i(v_i), b_{\text{-}i})] \geq \mathop{\mathrm{E}}_{\substack{v_{\text{-}i} \sim \mathcal{F} | v_i \\ b_{\text{-}i} = \sigma_{\text{-}i}(v_{\text{-}i})}} [v_i x_i(\sigma_i'(v_i), b_{\text{-}i}) - p_i(\sigma_i'(v_i), b_{\text{-}i})]$$

A mechanism is BIC if truth-telling, i.e. $\sigma_i(v_i) = v_i \forall i$*, is a Bayesian Nash Equilibrium.*

We start by assuming agents have independent value distributions $\mathcal{F}_i$. Equivalently, the joint distribution $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2 \times \cdots \times \mathcal{F}_n$. Fix $\sigma_{\text{-}i}$, define

$$x_i(b) = \mathop{\mathrm{E}}_{\substack{v_{\text{-}i} \sim \mathcal{F} | v_i \\ b_{\text{-}i} = \sigma_{\text{-}i}(v_{\text{-}i})}} [x_i(b, b_{\text{-}i})];$$

$$p_i(b) = \mathop{\mathrm{E}}_{\substack{v_{\text{-}i} \sim \mathcal{F} | v_i \\ b_{\text{-}i} = \sigma_{\text{-}i}(v_{\text{-}i})}} [p_i(b, b_{\text{-}i})].$$

We now state the general case of Myerson's lemma:

Lemma 2.4.2 (Myerson)
A mechanism (x, p) for the single parameter (linear agent) setting is BIC iff $\forall i$:

1. *$x_i(v_i)$ is monotone non-decreasing in v_i;*
2. *Satisfies the payment identity $p_i(v_i) = v_i x_i(v_i) - \int_0^v x_i(t)dt + p_i(0)$.*

Similarly, a mechanism (x, p) for the single parameter (linear agent) setting is DSIC iff $\forall i, \mathbf{v_{\text{-}i}}$:

1. *$x_i(v_i, \mathbf{v_{\text{-}i}})$ is monotone non-decreasing in v_i;*
2. *Satisfies the payment identity $p_i(v_i, \mathbf{v_{\text{-}i}}) = v_i x_i(v_i, \mathbf{v_{\text{-}i}}) - \int_0^v x_i(t, \mathbf{v_{\text{-}i}})dt + p_i(0)$.*

Proof: Omitted. ■

Observe the following implications of Myerson's Lemma:

1. Revenue Equivalence

   Two BIC mechanisms that have the same allocation rules will always have the same expected revenue. For example, in the Vickrey auction the agent with the highest value will win. In the First Price sealed bid auction where every buyer's value is drawn from some some distribution and so they use symmetric strategies, the buyer with the highest value will win. Since both mechanisms are BIC, we immediately conclude that these two mechanisms will yield the same revenue.

2. Revenue Linearity

   Observe the payment identity: if we average the allocation rules of two mechanism, then the expected revenue gets averaged.

3. For the single parameter (linear agent) setting, the optimal BIC mechanism is deterministic and DSIC.

   Proof: We prove for the single agent setting. First notice that by Myerson's lemma, the allocation rule $x(\cdot)$ in any deterministic mechanism that is DSIC must be a step function. In other words, the mechanism is simply the "pricing" mechanism: charge price θ for allocating the item and 0 for not allocating. Denote the step function $x(\cdot)$ whose jump point is θ as τ_θ. Thus the revenue for such mechanism is

   $$\text{Revenue}(\theta) = \theta \Pr[v \geq \theta] = \theta(1 - F(\theta))$$

   where F is the cumulative distribution function. Notice that the optimal deterministic DSIC mechanism will, by definition, optimize Revenue(θ).

   Now for the optimal BIC mechanism, by Myerson's lemma again, the allocation rule $x(\cdot)$ must be monotone non-decreasing in v. By basic result in functional analysis, we can write any non-decreasing function $x : \mathbb{R} \to [0, 1]$ as a distribution of step functions. Namely,

   $$x(v) = \int_0^\infty x'(\theta)\tau_\theta(v)d\theta$$

By the revenue linearity,

$$\text{Revenue}(x) = \int_0^\infty x'(\theta)\,\text{Revenue}(\tau_\theta)d\theta$$

Finally observe that Revenue(x) reaches maximal when Revenue(τ_θ) reaches its maximal pricing, which is the exactly when the mechanism is optimal, deterministic and DSIC. ■