CS364B: Frontiers in Mechanism Design
Lecture #18: Multi-Parameter Revenue-Maximization*

Tim Roughgarden†

March 5, 2014

1 Review of Single-Parameter Revenue Maximization

With this lecture we commence the fifth and final part of the course. All previous lectures focused on the objective of maximizing the welfare $\sum_{i=1}^{n} v_i(S_i)$ of an auction. In these last three lectures, we study the objective of maximizing the revenue $\sum_{i=1}^{n} p_i$ of an auction. Of course, all of the auctions that we've studied to generate revenue, but only as a side effect of the quest for incentive-compatible welfare-maximization. In effect, are we switching perspectives from that of a non-profit-maximizing entity (like a government) to that of a monopolist. Alternatively, we're moving from a competitive market, where competition might prevent monopoly pricing, to a non-competitive market. This change in objective leads to a quite different theory. One thing in common between this part and previous ones is that major progress has been made just in the past few years, in particular in the theoretical computer science literature.

1.1 The Challenges of Revenue-Maximization

The first reason that we've be so obsessed with welfare-maximization is that it is a fundamental objective. Many real-world combinatorial auction designs, for example for wireless spectrum licenses, are guided explicitly by welfare-maximization concerns.

The second reason is pedagogical. Welfare-maximization is special and, the difficulties of the last 17 lectures notwithstanding, offers fewer conceptual challenges than revenue-maximization. For example, if complexity is not a concern, then the VCG mechanism offers an astonishing guarantee: the mechanism is dominant-strategy incentive compatible (DSIC) and maximizes the welfare for every valuation profile (assuming truthful reporting). In this sense, the VCG mechanism reduces welfare maximization with privately held valuations to

†Department of Computer Science, Stanford University, 462 Gates Building, 353 Serra Mall, Stanford, CA 94305. Email: tim@cs.stanford.edu.

the same problem with publicly known valuations — the strong DSIC guarantee comes "for free."

Revenue maximization is different. By definition, welfare is defined extrinsic to any particular mechanism; revenue is a property of the mechanism itself. As a result, even in a trivial scenario with a single buyer and a single item, there is no "optimal" mechanism. (To maximize welfare, one would just give the item away for free.) Is a take-it-or-leave-it offer at \$20 better than one at \$10? The answers depends on the bidder's valuation, which is unknown to the seller. The most common way of comparing the revenue of two different mechanisms — inevitably, one has higher revenue on some inputs, the other on other inputs — is to use a prior distribution. There is then an unequivocally "optimal" incentive-compatible mechanism — the one with the highest expected revenue with respect to the prior distribution. Of course, changing the prior generally changes the optimal mechanism (cf. welfare-maximization, where the VCG mechanism is always optimal).[1]

For example, adopting a prior makes the single-buyer single-item revenue-maximization problem very easy to think about. The direct-revelation incentive-compatible mechanisms correspond to probability distributions over take-it-or-leave-it-offers.[2] The optimal mechanism for a prior distribution F to make a take-it-over-leave it offer in

$$\operatorname*{argmax}_{r \geq 0} r \cdot (1 - F(r)),$$

since r is the revenue from a sale and $1 - F(r)$ is the probability of a sale. Recall that such a price is called a *monopoly price* for F. Changing the prior F changes only the monopoly price, and not the structure of the optimal mechanism.

1.2 Recap: Myerson's Optimal Auction Theory

We saw last quarter that Myerson [1] gave a complete solution to Bayesian revenue-maximization in single-parameter environments.[3] For example, consider a single-item auction with n bidders, where bidder i's valuation is drawn independently from a distribution F_i. The optimal auction awards the item to the bidder with the highest "virtual valuation" $\varphi_i(v_i)$, where

$$\varphi_i(z) = z - \frac{1 - F_i(z)}{f_i(z)},$$

[1]In CS364A, Lectures #5 and #6, we used a prior distribution for exactly this purpose. We only studied DSIC mechanisms, so the bidders did not need to know the prior distribution. Here, we study Bayesian incentive-compatible (BIC) mechanisms as well, so we use the prior also to model how bidders reason about what action to take.

[2]With only one buyer, there is no need to differentiate between DSIC and BIC mechanisms. Since there are no other valuations to average over, the two concepts coincide.

[3]Recall that a *single-parameter environment* with n bidders can be represented as a feasible set X of n-vectors, with bidder i's welfare in outcome $\mathbf{x}$ given by $v_i \cdot x_i$ for a private parameter v_i. For example, a single-item auction environment can be represented by the n standard basis n-vectors (designating the winner) and the all-zero vector (designating no winner).

with no sale of all bidders' virtual valuations are negative.[4] For example, if F_i is uniform on $[0,1]$, then $\varphi_i(z) = 2z - 1$, which is negative on $[0, \frac{1}{2})$ and positive on $(\frac{1}{2}, 1]$ Payments are uniquely determined by "Myerson's Lemma" (see CS364A, Lecture #3). When bidders are i.i.d. and regular, meaning all F_i's are the same and regular, then the optimal auction is particularly simply. With identical and strictly increasing virtual valuation functions, the optimal auction awards the item to the highest bidder (unless all virtual valuations are nonpositive, leading to no sale). This auction is equivalent to the Vickrey auction with a reserve price of $\varphi^{-1}(0)$ — informally, an eBay auction with a judiciously chosen opening bid.

In general, when bidders are single-parameter with regular valuation distributions, the optimal auction simple maximizes the virtual welfare:

1. For each bidder i with report b_i, use the publicly known distribution F_i to compute its virtual bid $\varphi_i(b_i)$.

2. Choose the feasible outcome $\mathbf{x} \in X$ that maximizes $\sum_{i=1}^{n} \varphi_i(\mathbf{b}) x_i$, the virtual welfare according to the reported valuations.

3. Charge payments according to Myerson's Lemma to ensure that the mechanism is DSIC.

In other words, revenue-maximization in single-parameter environments reduces to virtual welfare-maximization.

The first reason this reduction is interesting is conceptual: it gives us a solid handle on what optimal mechanisms look like. A remarkable corollary of this reduction is that, even if we optimize over the space of all BIC mechanisms, there is an optimal mechanism that satisfies the much stronger DSIC guarantee.[5] Equally remarkably, with regular valuation distributions, optimizing over the space of randomized mechanisms yields an optimal mechanism that is deterministic.

The second reason to care about Myerson's reduction is computational. It implies that in every problem domain where welfare maximization is computationally tractable, revenue maximization is equally tractable — to solve the latter, one just forms the virtual valuations and feeds them into your favorite welfare-maximization algorithm.

In summary, for single-parameter problems, there is a very satisfying theory of revenue-maximizing auction. But the problems studied in this class, combinatorial auctions, are generally not single-parameter problems. *The goal of this part of the course is to understand the extent to which a "Myerson-like" theory of revenue-maximizing auctions is possible for multi-parameter problems.* We've seen some remarkable progress on this goal over the past couple of years, more than 30 years after Myerson's paper [1]. The goal of this lecture is to

[4] The fine print: we're assuming that each distribution has positive density on its support. We're also assuming that each distribution F_i is *regular*, which means that the corresponding virtual valuation function is strictly increasing. Myerson's optimal auction theory holds more generally for irregular distributions, although it involves additional details that we won't discuss here.

[5] We didn't actually prove this last quarter because we never discussed BIC mechanisms. Just as it's not difficult to extend Myerson's lemma and its proof to interim allocation rules (Lecture #12), it's not difficult to extend Myerson's optimality of virtual welfare maximization to the space of BIC mechanisms.

get a feel for the challenges involved. In the next two lectures we'll sink our teeth into the latest results.

2 Examples and Challenges in Multi-Parameter Revenue-Maximization

A single-buyer scenario already indicates the much greater complexity of optimal revenue-maximization in multi-parameter settings. Consider;

1. A single buyer and $m = 2$ non-identical items.
2. A prior distribution over additive valuations (v_1, v_2). Moreover, assume that v_i, v_2 are i.i.d. draws from a distribution F.

Have we simplified the problem to the point that it's trivial? After all, we know how to do revenue-maximization with one buyer and one item, and since the buyer's valuation is additive over the goods the obvious thing to do is to sell each item separately (at the monopoly price for F).

Example 2.1 (Bundling can be better) Consider the simplest imaginable distribution where each of v_1, v_2 is equally like to be 1 or 2 (independently). Selling the items separately yields an expected revenue of 2 (the price can 1 or 2, it doesn't matter). Another, equally legitimate incentive-compatible mechanism *bundles* the two items together and offers a take-it-or-leave-it offer for them. If prices at 3, this mechanism yields an expected revenue of 9/4 — the buyer buys the bundle except in the 1 in 4 chance that $v_1 = v_2 = 1$ — and is thus better than selling the items separately!

It should be clear that Example 2.1 is not at all pathological. Indeed, the lesson that bundling is better than selling items separately is particularly clear in the i.i.d. case as the number of items m goes to infinity: for large enough m the sum of the buyer's valuations for all items becomes highly concentrated, so offering only the "grand bundle" at a price slightly below the buyer's value for it has expected revenue almost equal to the expected welfare — an upper bound on the expected revenue of any mechanism, and well more than one can typically be achieved by selling items separately.

Perhaps bundling the items together is always optimal, at least in simple scenario that we're currently discussing?

Example 2.2 (Better than bundling) Consider a distribution F that is equally likely to take on the values 0, 1, or 2. Selling items separately at a price of 1 or 2 yields expected revenue 4/3. Selling the only as a bundle also yields expected revenue 4/3 (at a price of 2, with selling probability 2/3).

On the other hand, suppose we offer any one item at a price of 2, of both items together at the discounted price of 3. The buyer chooses its utility maximizing bundle given these

$v_1 \backslash v_2$	0	1	2
0	0	0	2
1	0	0	3
2	2	3	3

Table 1: The revenue matrix of Example 2.2

.

choices and prices.[6] To compute the expected revenue of this mechanism, consider the matrix in Table 1. It is obvious that buyer will buy nothing if $v_1, v_2 \leq 1$ — all three options yield negative utility. If the buyer has value 2 for 1 item and 0 for the other, than it purchases the item it wants at a price of 2. If is has value 2 for one item and at least 1 for the other, then it purchases the bundle of both items at a price of 3.[7] The resulting expected revenue is $\frac{3}{9} \cdot 3 + \frac{2}{9} \cdot 2 = \frac{13}{9} > \frac{4}{3}$, which is better than either selling items separately or selling only the grand bundle.

What happens in the next example is even weirder.

Example 2.3 (Randomization can be necessary) Suppose now that each of v_1, v_2 is drawn i.i.d. from a distribution F that 1 with probability $\frac{1}{6}$, 2 with probability $\frac{1}{2}$, and 4 with probability $\frac{1}{3}$. Consider the mechanism where the buyer has three options: pay 1 for a lottery ticket that yields a 50% probability of winning the first item; pay 1 for a lottery ticket that yields a 50% probability of winning the second item; or pay 4 to get both items with 100% probability. As usual, the buyer chooses its utility-maximizing bundle, and we can assume that ties are broken in favor of the highest-price bundle. We leave the following assertions as exercises:

(a) the auction above has expected revenue $3\frac{17}{36}$;

(b) every deterministic auction — where every outcome awards either nothing, the first item, the second item, or both items — has strictly less expected revenue.

In particular, selling a lottery ticket with a 50% probability of a realization at price 1 is not the same as selling an item at price 2. Prices affect the buyer's utility additively; the probability of item realization affects the buyer's value multiplicatively.

The upshot of these examples is that revenue-maximizing auctions are much more complicated in multi-parameter settings, even with just one buyer with an additive valuation over two items, than in single-parameter settings. The format of the optimal auction varies significantly with the prior F and, even with a very simple prior distribution, the optimal auction

[6]In the direct-revelation version of this mechanism, the buyer reports its valuation, and the auction picks the utility-maximizing choice on behalf of the buyer.

[7]We are assuming that ties are broken in favor of the seller. This is without loss of generality, in the sense that the seller can perturb prices by an arbitrarily small amount (e.g., from 2 to $2 - \epsilon$ and from 3 to $3 - 2\epsilon$) to ensure the desired tie-breaking.

need not be deterministic. This is sharp contrast with the single-parameter case, where with a single-buyer the optimal auction is just a suitably (deterministic) take-it-or-leave-it offer.

3 A Linear Programming Approach

How can we make sense of the complex dependence between the optimal auction and the prior distribution, even in the very simple setups in Examples 2.1–2.3? Closed-form explicit formulas, analogous to single-parameter virtual valuations, seem too simple to hope for. The problem would seem to require more powerful machinery, and we turn to linear programming to help us.

3.1 The Single-Buyer Case

A straightforward observation is that the revenue-maximizing mechanism with respect to a prior can be characterized as the solution to a linear program. This is not very explicit, but we've got to start somewhere.

Here is the new setup:

- a single buyer;
- a set U of m items;
- a finite set $V = \{v^1, \ldots, v^r\}$ of possible additive valuations (each an m-vector);
- a probability mass $f(v)$ for each $v \in V$; the $f(v)$'s are nonnegative and sum to 1.

Additive valuations are interesting to start with — recall Examples 2.1–2.3 — and we'll generalize them later. We will not need to assume that the buyer's values v_j and v_k for different items are independent.

The Revelation Principle (CS364A, Lecture #4) holds in the present setting, so we can restrict attention to direct-revelation mechanisms where truthful reports are always utility-maximizing. Then, we can formulate the following linear program. There is variable $x_j(v)$ for each $j \in U$ and $v \in V$, indicating the probability that the buyer receives the item j when it reports the valuation v. For example, in a mechanism that only offers the grand bundle at some price r, $x_j(v) = 1$ for all j when $\sum_{j \in U} v_j \geq r$ and $x_j(v) = 0$ for all j when $\sum_{j \in U} v_j < r$. The $x_j(v)$'s are 0 or 1 in a deterministic mechanism, but Example 2.3 shows that it's essential that we allow randomized allocation rules. There is also a decision variable $p(v)$ for each $v \in V$, indicating the payment from the buyer to the mechanism when its reports the valuation v.[8]

[8]The mechanism can use a randomized payment rule, but with a risk-neutral buyer and seller the only relevant quantity is the expected payment $p(v)$ when the buyer reports v. Also, there is no need to keep track of a separate payment for each item.

The objective function is to maximize revenue, which can be written as

$$\max \sum_{v \in V} f(v)p(v). \tag{1}$$

Since the $f(v)$'s are given, this is a linear objective function. The incentive-compatibility (IC) constraints, asserting that truthful reporting is utility-maximizing, can be written as

$$\sum_{j \in U} v_j x_j(v) - p(v) \geq \sum_{j \in U} v_j x_j(v') - p(v') \tag{2}$$

for every $v, v' \in V$. Note we are using the additivity of the buyer's valuation. Since the v's are given, there are linear constraints. The individual rationality (IR) constraints, requiring non-negative utility for a truthful buyer, can be expressed as the linear constraints

$$\sum_{j \in U} v_j x_j(x) - p(v) \geq 0 \tag{3}$$

for every $v \in V$. Naturally, we also require that

$$0 \leq x_j(v) \leq 1 \tag{4}$$

for every $j \in U$ and $v \in V$.

There is a natural correspond between IC and IR direct-revelation mechanisms and feasible solutions to the linear program (1)–(4). Given such a mechanism, defining the $x_j(v)$'s and $p(v)$'s according to their intended semantics yields a feasible solution. Given a feasible solution, one can define a mechanism in the obvious way: given a report v, allocate each item $j \in U$ to the buyer with probability $x_j(v)$, and charge a payment of $p(v)$.[9] This mechanism is well defined by (4), and is IC (by (2)) and IR (by (3)). This expected revenue of the mechanism and the objective function value of the corresponding feasible solution coincide.

The number of variables and constraints in the linear program (1)–(4) is polynomial in the size $|V|$ of the valuation space. This is large in some cases, but it's hard to see how one could ever encode the IC constraints with a much smaller formulation.

3.2 Multiple Buyers

An advantage of the linear programming approach is it can be easily extended in many different ways. For example, a buyer with a budget constraint B can be modeled by adding the constraints $p(v) \leq B$) for every $v \in V$. We also easily extends from one many buyers, in the following setup:

- n bidders;
- a set U of m items;

[9]Items can be allocated independently, or according to any other distribution with the correct marginals — that's all that a risk-neutral buyer with an additive valuation cares about.

- for each $i = 1, 2, \ldots, n$, a finite set $V_i = \{v_i^1, \ldots, v_i^{r_i}\}$ of possible additive valuations (each an m-vector);
- for each $i = 1, 2, \ldots, n$, a probability mass $f_i(v)$ for each $v \in V_i$; the $f_i(v)$'s are nonnegative and sum to 1.

We are assuming that bidder's valuations are independent, although the following linear program extends easily to correlated valuation distributions as well.

Let $\mathcal{V} = V_1 \times \cdots \times V_n$ denote the set of possible valuation profiles. There is a variable $x_{ij}(\mathbf{v})$ for each bidder i, item j, and valuation profile $\mathbf{v} \in \mathcal{V}$, indicating the probability that i gets j when the reported valuations are $\mathbf{v}$. There is a variable $p_i(\mathbf{v})$ for each bidder i and profile $\mathbf{v}$, indicating i's expected payment when the reported valuations are $\mathbf{v}$. The objective is still to maximize revenue:

$$\sum_{\mathbf{v} \in \mathcal{V}} \mathbf{F}(\mathbf{v}) \sum_{i=1}^{n} p_i(\mathbf{v}). \tag{5}$$

With multiple bidders, there is a distinction between Bayesian incentive compatibility and dominant-strategy incentive compatibility. We are interesting in the former, and they can be encoded as follows:

$$\sum_{\mathbf{v}_{-i} \in \mathcal{V}_{-i}} \mathbf{F}_{-i}(\mathbf{v}_{-i}) \left(\sum_{j \in U} v_{ij} x_{ij}(\mathbf{v}) - p_i(\mathbf{v}) \right) \geq \sum_{\mathbf{v}_{-i} \in \mathcal{V}_{-i}} \mathbf{F}_{-i}(\mathbf{v}_{-i}) \left(\sum_{j \in U} v_{ij} x_{ij}(v_i', \mathbf{v}_{-i}) - p_i(v_i', \mathbf{v}_{-i}) \right) \tag{6}$$

for every bidder i, true valuation v_i, and reported valuation v_i'.[10]

Similarly, the IR constraints – in this form usually called interim individual rationality (IIR) — can be written as

$$\sum_{\mathbf{v}_{-i} \in \mathcal{V}_{-i}} \mathbf{F}_{-i}(\mathbf{v}_{-i}) \left(\sum_{j \in U} v_{ij} x_{ij}(\mathbf{v}) - p_i(\mathbf{v}) \right) \geq 0 \tag{7}$$

for all i and $v_i \in V_i$. Finally, since each item is allocated to at most one bidder, we have the following feasibility constraints:

$$\sum_{i=1}^{n} x_{ij}(\mathbf{v}) \leq 1 \tag{8}$$

for every $j \in U$ and $\mathbf{v} \in \mathcal{V}$. (And of course, all the $x_{ij}(\mathbf{v})$'s should be nonnegative.)

There is again a natural correspondence between the feasible solutions of the linear program (5)–(8) and the BIC and IIR direct-revelation mechanisms. A feasible solution to (5)–(8) indicates what a mechanism should do: given reported valuations $\mathbf{v}$, assign each item j to a bidder i with probability $x_{ij}(\mathbf{v})$, and charge each bidder i a payment of $p_i(\mathbf{v})$. This

[10] With a non-product distribution $\mathbf{F}$, the only difference would be to average over $\mathbf{v}_{-i}$ with respect to the conditional distribution on $\mathbf{v}_{-i}$ given v_i.

The DSIC constraints corresponding to (6) would quantify over all possible reported valuations $\mathbf{v}_{-i}$, rather than averaging over the (truthfully reported) valuations $\mathbf{v}_{-i}$.

mechanism is well defined by (8) — each item j can assigned independently, for example — and is BIC and IIR by (6) and (7).

While the linear program (1)–(4) has size polynomial in the number $|V|$ of valuations of the single buyer, the linear program (5)–(8) has size polynomial in the number $|\mathcal{V}|$ of valuation *profiles*. The size of $\mathcal{V}$ scales exponentially with with the number of bidders n. This linear program is useless from a computational perspective for all but very small problems. In the next lecture, we show how to capture the set of BIC and IIR mechanisms with a much smaller set of variables, with size scaling polynomially in n. In addition to being more meaningful computationally, the quest for a smaller linear program will yield conceptually insights about the structure of revenue-maximizing mechanisms in multi-parameter settings.

References

[1] R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.