CSCI 5520: Foundations of Data Privacy **Lecture 9**
The Chinese University of Hong Kong, Spring 2015 10 March 2015

We consider the following scenario from game theory: n selfish players want to select a common, public outcome a among a set A of alternatives. Each player i has a private *valuation* function $v_i\colon A \to \mathbb{R}$ which assigns a numerical value $v_i(a)$ to every possible alternative $a \in A$. When A is finite, the valuation funtion can be described as a vector of values $(v_i(a))_{a\in A}$ in $\mathbb{R}^A$.

To make the selection, the players engage in a *game* $g\colon (\mathbb{R}^A)^n \to A$ that asks each player i for a *strategy* $s_i\colon A \to \mathbb{R}$ and outputs an outcome in A. The game is publicly known and is carried out by a party that is trusted to follow the correct instructions and not reveal any information apart from the outcome. The objective of each player is to maximize his valuation of the outcome: To do so, he may choose to submit a preference s_i that is different from his true valuation v_i.

For example, in an election with n voters and two candidates Alice and Bob, each voter's valuation function describes how much he likes each candidate. For example, my valuation function might be $v_{\text{Andrej}}(\text{Alice}) = 50$ and $v_{\text{Andrej}}(\text{Bob}) = -100$. From an economist's point of view, Alice winning the election would make me feel as happy as if I had won 50 dollars, while Bob winning would make me feel as unhappy as if I had lost 100.

If the election rule is to select the candidate that is more valued by the majority of the voters

$$g(s_1, \dots, s_n) = \begin{cases} \text{Alice}, & \text{if } s_i(\text{Alice}) > s_i(\text{Bob}) \text{ for a majority of } i\text{'s} \\ \text{Bob}, & \text{if not} \end{cases}$$

then each voter has an incentive to provide his real valuation v_i as his strategy s_i: Regardless of what the other voters do can only make the election of his preferred candidate more likely. This is very desirable feature of the majority vote.

Definition 1. Strategy s_i^* is a *dominant strategy* for player i in game g if for every choice of strategies $s_1, \dots, s_n$,

$$u_i(g(s_1, \dots, s_{i-1}, s_i^*, s_{i+1}, \dots, s_n)) \geq u_i(g(s_1, \dots, s_{i-1}, s_i, s_{i+1}, \dots, s_n)).$$

The strategy $s_i = v_i$ is called *truthful*. A game g is *dominant strategy truthful* if for each player the truthful strategy is a dominant strategy.

Unfortunately few interesting games are dominant strategy truthful. An example of one that is not is a three candidate election decided by plurality vote (you can work out why). One of our examples will achieve a relaxed notion of truthfulness. The other one will assume a more general model of a game in which payments can be made to "buy" utility.

Game-theoretic mechanism design is concerned with the design of games whose outcome meets certain requirements. For example, we might want to design an election that guarantees some aggregate happiness among the voters.

1 Digital goods auctions

In a digital goods auction, I have an unlimited number of copies of a product I would like to sell, for example a digital photograph or a piece of software. The alternatives are the possible prices p I can set for my item; each price is a number in the interval $[0, 1]$.

Player i's valuation describes how much he would value owning the item if it was sold at price x. Let us assume for simplicity (as it won't make difference) that player i's violation is completely described by a secret cutoff value $v_i \in [0, 1]$ so that

$$v_i(x) = \begin{cases} v_i - x, & \text{if } x \leq v_i \\ 0, & \text{if } x > v_i. \end{cases}$$

The objective of a digital goods auction is to maximize the revenue from all the players

$$f(v, p) = p \cdot (\text{number of players that buy at price } p) = p \cdot |\{i \colon v_i \geq p\}|.$$

For this purpose, I solicit the utilities from all the players and reveal a price for the product. If my mechanism outputs the value of p that maximizes $f(s, p)$ then an player has an incentive to bid a value s_i smaller than v_i.

Unfortunately no mechanism (except one that outputs a fixed price disregarding its inputs) of this form is dominant strategy truthful. Differential privacy can help us achieve an approximate notion of truthfulness.

Definition 2. A strategy s_i^* is *ε-approximately dominant* for player i in game g if for every choice of strategies $s_1, \dots, s_n$,

$$v_i(g(s_1, \dots, s_{i-1}, s_i^*, s_{i+1}, \dots, s_n)) \geq e^{-\varepsilon} v_i(g(s_1, \dots, s_{i-1}, s_i, s_{i+1}, \dots, s_n)).$$

A game is ε-approximately dominant strategy truthful if being truthful is an ε-approximately dominant strategy for every player.

This definition gives some incentive to players to misreport their utilities, but they can never gain more than an ε-fraction of utility by doing so. One interpretation is that if I value telling the truth by an amount of about ε then being truthful is a (truly) dominant strategy.

To understand why optimizing f is not approximately dominant strategy truthful, suppose half the other players have valuation 1 and the other half have valuation 1/2. Then my effect of on the optimal value of f is significant. Without my presence, the prices $p = 1/2$ and $p = 1$ would yield the same revenue, so my bid essentially determines the price that is chosen. However, my bid has little effect on the overall revenue. If we can ensure that no single bid significantly influences the outcome then I wouldn't have much of an incentive to underbid.

We now describe a mechanism that achieves approximate dominant strategy truthfulness and outputs an approximately optimal price with high probability. We will set the parameters ε and K later.

```
Mechanism $A$: On input $s = (s_1, \ldots, s_n) \in [0,1]^n$,
    For every integer $k \in [K]$ calculate $p_i = \exp(\varepsilon f(s, k/K))$.
    Output the price $k/K$ with probability proportional to $p_i$.
```

In other words, the price is chosen by applying the exponential mechanism with utility function f. The function f is 1-Lipschitz in the inputs $s_1, \ldots, s_n$ (for any value of $p = k/K$), so by Theorem 5 from Lecture 2, so mechanism A is ε-differentially private.

Theorem 3. *If a mechanism M is ε-differentially private then any strategy for any player is ε-approximately dominant in expectation over the randomness of the mechanism.*

In particular, being truthful is an ε-approximately dominant strategy.

Proof. Let s_i^*, s_i be any pair of strategies for player i. We will write (s_i, s_{-i}) as a shorthand for $(s_1, \ldots, s_i, \ldots, s_n)$.

$$\begin{aligned} \mathrm{E}[v_i(M(s_i, s_{-i}))] &= \sum_a v_i(a) \Pr[M(s_i, s_{-i}) = a] \\ &\le \sum_a v_i(a) e^{\varepsilon} \Pr[M(s_i^*, s_{-i}) = a] \\ &= e^{\varepsilon} \mathrm{E}[v_i(M(s_i^*, s_{-i}))]. \end{aligned}$$

□

We now show that this mechanism achieves close to optimal revenue. Letting ℓ denote the loss in revenue we are willing to tolerate, we set $K = 2n/\ell$ and show that

Theorem 4. *Let p^* be the price that maximizes the revenue $f(v, p)$. Then for every ℓ*

$$\Pr[f(v, A(v)) < f(v, p^*) - \ell] < \frac{2n}{\ell} \exp(-\varepsilon \ell).$$

For $\ell = (1/\varepsilon) \log n$, the probability that the mechanism loses more than ℓ units of revenue vanishes as n grows.

Proof. Let k be the largest integer so that $k/K \le p^*$. Then $f(k/K, v)$ is at least $f(p^*) - n/K$ because each player that buys at price p^* will also buy at price k/K, while selling at this lower price will incur a loss of at most $1/K$ of his revenue. By Theorem 6 from Lecture 2, the probability that A produces an outcome of utility smaller than $f(k/K, v) - t$ is less than $K \exp(-\varepsilon t/2)$. Setting $t = n/K$ proves the theorem. □

2 Payments and social welfare

A mechanism with payments produces, in addition to an outcome, a payment $p_i \in \mathbb{R}$ to be billed to player i.

Suppose I want to give away a painting; my objective is not to make money but to award it to the player who values it the most. Player i has a valuation $v_i > 0$ for the painting. The player's utility equals $v_i - p_i$ if he gets the painting at price p_i and zero if he doesn't get the painting.

One possibility is to ask every player to report their valuation and award the painting for free to the player that makes the highest bid. This is clearly not dominant strategy truthful as players have an incentive to overreport. Another possibility is to give the painting to the player with the highest bid and charge him the amount he bid. Then a player may have an incentive to underreport if he thinks he can get the painting for less.

The Vickrey auction mechanism awards the painting to the player that submits the highest bid and charges him the amount of the second highest bid. This mechanism is dominant strategy truthful by a case analysis. Take any player i and consider an arbitrary strategy for the other players; let b be their largest among their bids.

- If $b > v_i$ then player i's overall utility from winning would be negative, so bidding v_i and not winning can only improve his utility;
- If $b \leq v_i$ then bidding any value above b will bring player i same overall utility as bidding v_i; bidding any value below b will bring him zero utility, while a truthful bid of v_i brings him positive utility.

Publishing the winner of an auction entails a loss of privacy which may affect the player's valuation of the outcome. For example, although I really value the painting, it would be quite embarrassing for me if it was discovered how much I spent on it. I could set my valuation lower to reflect the cost of this embarrassment. However, if privacy considerations were taken into account in the design of the mechanism then my embarrassment would be avoided (or at least mitigated) resulting in higher utility. As an extreme example, if the mechanism was a lottery which assigns the painting to a random player at no cost then there would be no embarrassment at all in winning the painting.

In general it may be desirable to strike a balance between achieving a socially valuable objective (giving the painting away to the player who values it the most) and preserving the privacy of the players. To explain we first extend the Vickrey auction to a setting with arbitrary outcomes where the objective is to maximize social welfare.

The Vickrey-Clarke-Groves mechanism

Let A be a set of alternatives. An n-player *mechanism with payments* is a game $g \colon (\mathbb{R}^A)^n \to A$ together with *payment functions* $p_1, \dots, p_n \colon (\mathbb{R}^A)^n \to \mathbb{R}$.

In the Vickrey auction the alternatives are $A = \{1, \dots, n\}$ where alternative i means "player i gets the painting", $g(s_1, \dots, s_n)$ chooses an i that maximizes s_i, p_i is the second largest value among $s_1, \dots, s_n$, and $p_j = 0$ for $j \neq i$.

For a collection of valuations $v_1, \dots, v_n$ and an outcome a the *social welfare* is the value $v_1(a) + \dots + v_n(a)$. In the Vickrey auction, social welfare is maximized by giving the painting to the player with the highest valuation.

For a game with payments, dominant strategy truthfulness means that the *utility* I extract after making the payment is highest for the truthful strategy, regardless of other players' strategies.

Definition 5. A strategy s_i^* is *dominant* for player i in a mechanism with payments if for every choice of strategies $s_1, \dots, s_n$,

$$v_i(g(s_i^*, s_{-i})) - p_i(s_i^*, s_{-i}) \geq v_i(g(s_i, s_{-i})) - p_i(s_i, s_{-i}).$$

The mechanism is *dominant strategy truthful* if for every player the truthful strategy is a dominant strategy.

The Vickrey-Clarke-Groves (VCG) mechanism is a dominant strategy truthful mechanism that chooses the outcome with maximum social welfare. It has two additional desirable features: payments are always positive (I cannot "make money" just by participating) and utilities are always positive (I cannot lose value by participating).

```
Mechanism $VCG$: On input $s_1, \dots, s_n \in \mathbb{R}^A$,
    Output an outcome $a^* \in A$ that maximizes the social welfare $\sum_i s_i(a^*)$.
    Charge player $i$ a payment of $p_i = \max_{a \in A} \sum_{j \neq i} s_j(a) - \sum_{j \neq i} s_j(a^*)$.
```

In words, the payment of player i is the maximum social welfare if player i didn't participate minus the social welfare of the others when player i is present; this value is nonnegative and always smaller than player's reported valuation $s_i(a^*)$.

Theorem 6. *Mechanism VCG is dominant strategy truthful.*

Proof. More generally, we will show that for any choice of functions $h_{-1}, \dots, h_{-n}$, where h_{-i} does not depend on s_i, a payment of $p_i(s) = h_{-i} - \sum_{j \neq i} s_j(a^*)$ to player i is dominant strategy truthful. Player i's utility is

$$v_i(a^*) + \sum_{j \neq i} s_j(a^*) - h_{-i}(s) \leq \max_{a \in A}\Big(v_i(a) + \sum\nolimits_{j \neq i} s_j(a)\Big) - h_{-i}(s)$$

The maximum on the right hand side is attained for $a = VCG(v_i, s_{-i})$, namely when player i's strategy is v_i. □

3 Private optimization of social welfare

The VCG mechanism is deterministic so it is clearly not private. A natural idea towards privacy is to randomize the choice of outcome, but give more weight to outcomes that achieve higher social welfare. This brings to mind the exponential mechanism:

```
Mechanism $PrivateVCG$: On input $s_1, \dots, s_n \in \mathbb{R}^A$,
    Output outcome $a \in A$ with probability $D(a) = \frac{1}{Z}\exp\big((\varepsilon/2)\sum_i s_i(a)\big)$.
    For each player $i$,
        Let $D_{-i}$ be the distribution that assigns probability $\frac{1}{Z_{-i}}\exp\big((\varepsilon/2)\sum_{j \neq i} s_j(a)\big)$ to outcome $a$.
        Charge player $i$ a payment of $p_i$ where
```

$$p_i = \mathrm{E}_{a\sim D_{-i}}\Big[\sum\nolimits_{j\neq i} s_j(a)\Big] - \mathrm{E}_{a\sim D}\Big[\sum\nolimits_{j\neq i} s_j(a)\Big] + \frac{2}{\varepsilon}(H(D_{-i}) - H(D)).$$

Here $H(D)$ is the natural entropy of the distribution D:

$$H(D) = \mathrm{E}_{a\sim D}[-\ln D(a)].$$

and Z and Z_{-i} are normalization constants so that the probabilities add to one. The curious choice of prices will come out naturally from the analysis.

The outcome of $PrivateVGC(s)$ is chosen from the exponential mechanism with utility function $Soc(s,a) = \sum_i s_i(a)$. Assuming the strategies are bounded (i.e. $s_i(a) \in [0,1]$), Soc is 1-sensitive so the mechanism is ε-differentially private and close to optimal in social welfare (if there are not too many alternatives):

$$\Pr[PrivateVCG(s) < VCG(s) - \ell] < |A|\exp(-\varepsilon\ell/2).$$

Theorem 7. *Mechanism* $PrivateVCG$ *is dominant strategy truthful in expectation. Moreover,* $0 \leq p_i \leq \mathrm{E}[s_i(PrivateVCG(s))]$ *for every player* i.

The proof of this theorem relies on the following lemma, which can be proved using calculus.

Lemma 8. *For any* $\lambda > 0$, *any finite set* A, *and any function* $F\colon A \to \mathbb{R}$, *the quantity* $\mathrm{E}_{a\sim D}[F(a)] + (1/\lambda)H(D)$ *is maximized by the distribution* D *over* A *that chooses outcome* $a \in A$ *with probability proportional to* $\exp(\lambda F(a))$.

Proof of Theorem 7. Let $\mathcal{A}$ be the set of probability distributions over the set A. Consider the VCG mechanism for choosing an alternative $D \in \mathcal{A}$ with $n+1$ players and the following valuations: The valuation of player $i \in \{1,\ldots,n\}$ is $\overline{v}_i(D) = \mathrm{E}_{a\sim D}[v_i(a)]$ and the valuation of player $n+1$ is $\overline{v}_{n+1}(D) = (2/\varepsilon)H(D)$.

The outcome of this mechanism is the distribution D that maximizes the social welfare

$$Soc(D) = \sum_{i=1}^{n+1} \overline{s}_i(D) = \mathrm{E}_{a\sim D}\Big[\sum\nolimits_{i=1}^{n} s_i(a)\Big] + \frac{2}{\varepsilon}H(D)$$

which by Lemma 8 is exactly the distribution $D(a) = \frac{1}{Z}\exp((\varepsilon/2)\sum_{i=1}^{n} s_i(a))$. Therefore the expected social welfare of the outcome of $PrivateVCG(s)$ equals the social welfare of the outcome of $VGC(\overline{s})$.

By a similar argument, the payment by player i in $PrivateVCG(s)$ equals the payment by player i in $VCG(\overline{s})$. By Theorem 6 the VCG mechanism is dominant strategy truthful, so $PrivateVCG$ is dominant strategy truthful in expectation. The "moreover" part follows from the fact that VCG payments are nonnegative and bounded by the players' reported valuations. □

One unsatisfying aspect of Theorem 7 is that it only gives nonnegative utility in expectation. I don't know if it is possible to strengthen this guarantee to hold with probability 1.

References

These notes are based on Chapter 10 of the survey *The Algorithmic Foundations of Differential Privacy* by Cynthia Dwork and Aaron Roth.

The presentation of the Vickrey-Clarke-Groves mechanism is based on these notes of Avrim Blum. A proof of Lemma 8 which dates back to Boltzmann can be found for example in these notes of David Aldous.