CMSC5706 Topics in Theoretical Computer Science

Week 8: Social Choice and Mechanism Design

Instructor: Shengyu Zhang

Social welfare

- Motivating example 1.
- Each year we interview and recruit graduate students.
- A panel of 4-6 professors attend the interview and give individual rank of the 20-30 candidates.
- We need to aggregate these rankings to get a final ranking for the department.
- *Question: How to aggregate rankings?*

Social choice

- Motivating example 2.
- A small number of candidates run for president.
- A large number of voters, each gives a ranking of the candidates

- *Question: Who should win?*

Formal setting

- A: set of alternatives/candidates.
- I: set of n voters/professors.
- L: set of linear orders of A.
  - A linear order is a full ranking of alternatives in A.
  - Equivalently, a permutation of alternatives in A.
  - E.g. $a_4 \prec a_3 \prec a_1 \prec a_5 \prec a_2$ for $A = \{a_1, a_2, a_3, a_4, a_5\}$
- Each voter i has a linear order $\prec_i \in L$.

Formal setting

- A: set of alternatives/candidates.
- I: set of n voters/professors.
- L: set of linear orders of A.
- Each voter i has a linear order $\prec_i \in L$.
- Social welfare function: a function $F: L^n \to L$.
- Social choice function: a function $f: L^n \to A$.

Social welfare

- Let's consider social welfare functions first.
- What would be a good social welfare function F?

Desirable properties

- Unanimity: For every $\prec \in L$, $F(\prec, \ldots, \prec) = \prec$.
  - If everyone has the same preference list $\prec$, then we should just use that.

Desirable properties

- Independence of irrelevant alternatives: $\forall a, b \in A$, $\forall \prec_1, \dots, \prec_n, \prec'_1, \dots, \prec'_n \in L$, let $\prec = F(\prec_1, \dots, \prec_n)$ and $\prec' = F(\prec'_1, \dots, \prec'_n)$. Then

  $a \prec_i b \Leftrightarrow a \prec'_i b, \forall i$ implies $a \prec b \Leftrightarrow a \prec' b$.

  - The social preference between any a and b depends only on the voters' preferences between a and b.
  - If each voter i changes his ranking from $\prec_i$ to $\prec'_i$, as long as they each don't change the relative preference between a and b, then they won't change the final comparison between a and b.

Impossibility 1

- *Arrow's theorem*. *If* $|A| \geq 3$, *then only dictatorship satisfies both unanimity and independence of irrelevant alternatives.*
- A dictatorship is a social welfare function $F(\prec_1, \ldots, \prec_n) = \prec_i$ for some $i \in [n]$.
  - It's not a voting any more.
- Arrow's theorem says that *there is no good social welfare function*.

Social choice?

- Social choice needs to get only one winner.
  - Easier task than social welfare.

- *Question: Is there a good social choice function?*

Condorcet's Paradox

- Consider an election with two candidates and n voters.
- Majority is a good idea: Whoever gets more votes wins.
- What about three candidates?
- One idea: Use pairwise comparisons.
- But this runs into a problem.

Condorcet's Paradox

- Consider 3 voters with the following preferences for the three candidates a, b, c.
  - $a \prec_1 b \prec_1 c$
  - $b \prec_2 c \prec_2 a$
  - $c \prec_3 a \prec_3 b$
- Between (a, b): voter 1 and voter 3 prefer b.
- Between (b, c): voter 1 and voter 2 prefer c.
- Between (c, a): voter 2 and voter 3 prefer a.

Condorcet's Paradox

- Between (a, b): voter 1 and voter 3 prefer b.
- Between (b, c): voter 1 and voter 2 prefer c.
- Between (c, a): voter 2 and voter 3 prefer a.
- If a is elected, voter 1 and 3 would say "*Hey*, *why not a better candidate b*"?
  - More people (2 out of 3) prefer b to a, why should a win?
- If b or c is elected, similar issue appears as well.
- This is called Condorcet's Paradox.

Desirable properties

- Back to Question: Is there a good social choice function?
- A function is bad if it can be strategically manipulated: For some $\prec_1, \ldots, \prec_n \in L$ and some $\prec_i' \in L$, we have that $a \prec_i a'$ where $a = f(\prec_1, \ldots, \prec_n)$ and $a' = f(\prec_1, \ldots, \prec_i', \ldots, \prec_n)$.
  - You can change the final outcome from a to some a' who you like more (according to your real preference $\prec$), by presenting a fake preference list $\prec'$.
- A function f is called incentive compatible if it cannot be manipulated.

An equivalent view

- A social choice function f is monotone if different $a = f(\prec_1, \ldots, \prec_n)$ and $a' = f(\prec_1, \ldots, \prec_i', \ldots, \prec_n)$ implies $a' \prec_i a$ and $a \prec_i' a'$.
  - If your real preference is $\prec_i$, then faking it to $\prec_i'$ would only make the final outcome worse.
  - Same if your real preference is $\prec_i'$.
- incentive compatible $\Leftrightarrow$ monotone.

Impossibility 2

- Voter i is a dictator is f always outputs whoever ranks the highest in $\prec_i$.
- f is a dictatorship if some voter i is a dictator.
  - Again, dictatorship is not a good voting function.
- *Gibbard-Satterthwaite Theorem. If $|A| \geq 3$, then any incentive compatible social choice function f onto A is a dictatorship.*
- “You can’t ask for both.”

Mechanisms with money

- So far we've seen that there is no good social welfare/choice function.
- One way to get out of this dilemma is to use money.
- The preference of player i is given by a valuation function $v_i: A \to \mathbb{R}$.
  - $v_i(a)$ is the value that Player i assigns to alternative a.
- If a is chosen and Player i is additionally given some quantity m of money, then Player i's utility is $u_i = v_i(a) + m$.

A simple auction

- 1 item, n players.
- Player i has a value w_i that he is willing to pay for this item.
- If Player i gets the item at price p, then his utility is $w_i - p$.
- This is a social choice problem.
  - $A = \{\text{candidate } i \text{ wins: } i \in I\}$
  - Valuation: $v_i(i \text{ wins}) = w_i$, and $v_i(j \text{ wins}) = 0$, $\forall j \neq i$.

Who gets the item

- Question 1: Who gets the item?

- Answer: whoever values it the most.
  - Namely, $i \in \operatorname{argmax}_j w_j$.

- Question 2: Pays how much?

Two natural payments

- **No payment**. Give the item for free to a player with the highest w_i.
- Issue: Player i will manipulate this by exaggerating his w_i.
- **Pay your bid**. The winner i pays the declared bid w_i.
- Issue: His utility becomes $w_i - w_i = 0$.
- Thus he has incentive to declaring a lower value $w_i' < w_i$ with the hope that he still wins.
  - And his utility becomes $w_i - w_i' > 0$.

Vickrey's second price auction

- The winner is the player i with the highest declared value of w_i.
- And he pays the second highest declared bid $\max_{j \neq i} w_j$.
- *Theorem. For any* $w_1, \ldots, w_n$ *and any* w_i'*, let* u_i = *Player i's utility when bidding* w_i*, and* u_i' = *Player i's utility when bidding* w_i'*. Then* $u_i \geq u_i'$.
- The best strategy for each player is to report his real value, regardless of how others bid.
  - Even if others are cheating.

Vickrey's second price auction

- *Theorem*. *For any* $w_1, \ldots, w_n$ *and any* w_i', *let*
  u_i = *Player i's utility when bidding* w_i, *and*
  u_i' = *Player i's utility when bidding* w_i'.
  Then $u_i \geq u_i'$.
- Proof.
  Case 1. Player i wins by declaring w_i. Let p be the second highest reported value. Then $u_i = w_i - p$. For any attempted manipulation w_i':
  - $w_i' \geq p$: Player i still wins, and still pays p. So $u_i' = u_i$.
  - $w_i' < p$: Player i loses and gets payoff $0 \leq w_i - p = u_i$.

Vickrey's second price auction

- *Theorem*. *For any* $w_1, \dots, w_n$ *and any* w_i', *let*
  u_i = *Player i's utility when bidding* w_i, *and*
  u_i' = *Player i's utility when bidding* w_i'.
  Then $u_i \geq u_i'$.
- Case 2. Player i loses by declaring w_i. Then $u_i = 0$. The winner j has $w_j \geq w_i$. For any attempted manipulation w_i':
  - $w_i' < w_j$: Player i still loses, and get the same payoff 0.
  - $w_i' \geq w_j$: Player i wins and needs to pay w_j, so his payoff is $u_i' = w_i - w_j \leq 0 = u_i$.

- This mechanism is simple but elegant.
- It computes an $argmax$ function of n private numbers.
- It's like Adam Smith's *invisible hand*: despite private information and pure selfish behavior, social welfare is achieved.

Formal treatment of mechanism

- Each player i has a valuation function $v_i: A \to \mathbb{R}$, where $v_i \in V_i$.
- $V_i \subseteq \mathbb{R}^A$ is a commonly known set of all possible valuation functions for player i.
- The complete social choice has two parts
  - Alternative chosen
  - Transfer of money
- A mechanism is a social choice function $f: V_1 \times \cdots \times V_n \to A$ and a vector of payment functions $p_1, \ldots, p_n$, where $p_i: V_1 \times \cdots \times V_n \to \mathbb{R}$ is the amount that player i pays.

Truthfulness

- A mechanism $(f, p_1, \ldots, p_n)$ is *incentive compatible* if $\forall i \in [n], \forall v_1 \in V_1, \ldots, v_n \in V_n, \forall v_i' \in V_i$, $v_i(a) - p_i(v_i, v_{-i}) \geq v_i(a') - p_i(v_i', v_{-i})$, where $a = f(v_i, v_{-i})$ and $a' = f(v_i', v_{-i})$.
  - Player i would prefer "telling the truth" v_i to the mechanism rather than any possible "lie" v_i', since lying gives less utility.
- Such mechanism is also called *strategy-proof* or *truthful*.

VCG mechanism

- Social welfare of alternative $a \in A$: $\sum_i v_i(a)$
  - sum of valuations of all players for this alternative.
- A mechanism $(f, p_1, \dots, p_n)$ is a Vickrey-Clarke-Groves (VCG) mechanism if
  - $f(v_1, \dots, v_n) \in argmax_{a \in A} \sum_i v_i(a)$, i.e. f maximizes the social welfare, and
  - $p_i(v_1, \dots, v_n) = h_i(v_{-i}) - \sum_{j \neq i} v_j\big(f(v_1, \dots, v_n)\big)$ for some function $h_i: V_{-i} \to \mathbb{R}$

Intuition

- Note that the price that Player i needs to pay contains a term $-\sum_{j\neq i} v_j\big(f(v_1, \dots, v_n)\big)$.
- That is, he is paid $\sum_{j\neq i} v_j\big(f(v_1, \dots, v_n)\big)$.
- Plus his valuation $v_i(a)$ of getting a, he has $\sum_j v_j(a)$, the social welfare.
- Thus his payoff is social welfare minus $h_i(v_{-i})$, something unrelated to his v_i.
- So maximizing his own payoff is the same as maximizing the social welfare, which is achieved by reporting the true v_i.

Incentive compatible

- *Theorem. Every VCG mechanism is incentive compatible.*
- Proof. Need to show: for player i with valuation v_i, utility when declaring v_i is ≥ utility when declaring v_i'.
- Let $a = f(v_i, v_{-i}), a' = f(v_i', v_{-i})$.
- Utility u_i when declaring v_i: $v_i(a) + \sum_{j \neq i} v_j(a) - h_i(v_{-i})$.
- Utility u_i' when declaring v_i': $v_i(a') + \sum_{j \neq i} v_j(a') - h_i(v_{-i})$.
- Recall def of VCG: $a = f(v_1, \dots, v_n) \in argmax_{b \in A} \sum_i v_i(b)$
- Therefore $\sum_j v_j(a) \geq \sum_j v_j(a')$.
- Thus $u_i = \sum_j v_j(a) - h_i(v_{-i})$

$$\geq \sum_j v_j(a') - h_i(v_{-i})$$
$$= v_i(a') + \sum_{j \neq i} v_j(a') - h_i(v_{-i})$$
$$= u_i'.$$

What h_i to choose?

- $h_i = 0$: the mechanism pays the players.
- But usually the mechanism wants to get some money from the players.
- Clarke pivot rule: $h_i(v_{-i}) = \max_{b \in A} \sum_{j \neq i} v_j(b)$.
- The payment of player i is $p_i(v_1, \ldots, v_n) = \max_{b \in A} \sum_{j \neq i} v_j(b) - \sum_{j \neq i} v_j(a)$, where $a = f(v_1, \ldots, v_n)$.
- Intuitively, i pays the damage he causes---the difference between the social welfare of the others with and without his participation.

Properties

- A mechanism is individually rational if $v_i\big(f(v_1, \dots, v_n)\big) - p_i(v_1, \dots, v_n) \geq 0$.
- A mechanism has no positive transfers if $p_i(v_1, \dots, v_n) \geq 0$.
  - no player is paid money.
- *Theorem. A VCG mechanism with Clarke pivot payments makes no positive transfers. If $v_i(a) \geq 0, \forall v_i \in V_i$ and $a \in A$, then it is also individually rational.*

Properties

- *Theorem. A VCG mechanism with Clarke pivot payments makes no positive transfers. If $v_i(a) \geq 0, \forall v_i \in V_i$ and $a \in A$, then it is also individually rational.*
- Proof.
  individual rationality: the utility of i is
  $v_i(a) + \sum_{j \neq i} v_j(a) - \sum_{j \neq i} v_j(b)$
  $= \sum_j v_j(a) - \sum_{j \neq i} v_j(b)$
  $\geq \sum_j v_j(a) - \sum_j v_j(b) \quad (\because v_i(b) \geq 0)$
  $\geq 0 \quad (\because a$ maximizes $\sum_j v_i(a)$ in VCG)

Properties

- *Theorem. A VCG mechanism with Clarke pivot payments makes no positive transfers. If $v_i(a) \geq 0, \forall v_i \in V_i$ and $a \in A$, then it is also individually rational.*
- No positive transfer: $p_i(v_1, \ldots, v_n) = \sum_{j \neq i} v_j(b) - \sum_{j \neq i} v_j(a) \geq 0$, because b maximizes $\sum_{j \neq i} v_j(b)$ in Clarke pivot rule.

Back to single-item auction

- For single-item auction,
  VCG + Clarke pivot rule ⇒ 2nd price auction
- $A = \{P_1\ wins, P_2\ wins, \ldots, P_n\ wins\}$.
- $v_i(P_j\ wins) = \begin{cases} w_i & j = i \\ 0 & j \neq i \end{cases}$.
- $V_i = \{above\ v_i\ for\ any\ w_i \geq 0\}$
- $h_i(v_{-i})$ = the highest w_j among $j \neq i$
- $f(v_1, \ldots, v_n)$ is maximized by picking $P_i\ wins$ for an i with the largest w_i.
- $p_i = h_i(v_{-i}) - \sum_{j \neq i} v_j = \begin{cases} w_{j^*} & P_i\ wins \\ w_{j^*} - w_{j^*} = 0 & P_i\ doesn't\ win \end{cases}$
  - where j^* maximizes w_j among $j \neq i$.

Example 2 of VCG

- 1 buyer, n sellers.
- VCG Mechanism:
- The buyer gets the item from a seller with the lowest bid.
- The buyer pays to him only.
- The payment amount is the second lowest bid.
- Sometimes called "Reverse auction".

Example 3

- k identical items.
- n bidders, each interested in getting 1 item.
- VCG Mechanism:
- The k highest bidders get the k items (one for each).
- The i's highest bidder pays the $(i + 1)$'st highest offered price.