Mechanism design is about designing games in which some desired objective is achieved when all players play selfishly.

**Reverse game theory:** Rather than analyzing a given game, design one that fits your needs.

In this talk, we will see several mechanisms where the best strategy is to tell the truth.
Motivation

Public Project Problem

A public good is to be constructed, with cost $c$. There are $n$ agents who will enjoy this good, each with appreciation $v_i \geq 0$.

- If the project is accepted, everyone pays $c/n$; the utility of agent $i$ is $v_i - c/n$.
- If the project is rejected, no cost, and all utilities are 0.

**Decision rule:** Accept iff $\sum_{i=1}^{n} v_i \geq c$. 

Possible manipulation:
- If $v_i > c/n$, then Agent $i$ should declare $c$ and guarantee that the project is accepted!
- If $v_i < c/n$, and if everyone else is truthful, then Agent $i$ could try declaring 0.

Mechanism design: design games that selfish players cannot manipulate.
Motivation

Public Project Problem

A public good is to be constructed, with cost $c$. There are $n$ agents who will enjoy this good, each with appreciation $v_i \geq 0$.

- If the project is accepted, everyone pays $c/n$; the utility of agent $i$ is $v_i - c/n$.
- If the project is rejected, no cost, and all utilities are 0.

Decision rule: Accept iff $\sum_{i=1}^{n} v_i \geq c$.

Possible manipulation:
- If $v_i > c/n$, then Agent $i$ should declare $c$ and guarantee that the project is accepted!
- If $v_i < c/n$, and if everyone else is truthful, then Agent $i$ could try declaring 0.

Mechanism design: design games that selfish players cannot manipulate.
Motivation

Public Project Problem
A public good is to be constructed, with cost $c$. There are $n$ agents who will enjoy this good, each with appreciation $v_i \geq 0$.

- If the project is accepted, everyone pays $c/n$; the utility of agent $i$ is $v_i - c/n$.
- If the project is rejected, no cost, and all utilities are 0.

Decision rule: Accept iff $\sum_{i=1}^{n} v_i \geq c$.

Possible manipulation:
- If $v_i > c/n$, then Agent $i$ should declare $c$ and guarantee that the project is accepted!
- If $v_i < c/n$, and if everyone else is truthful, then Agent $i$ could try declaring 0.

Mechanism design: design games that selfish players cannot manipulate.
An **outcome** is to be chosen by a *central authority* from a set $A$. There are $n$ agents with different preferences.

0. Each agent $i$ has a true **valuation** $v_i : A \to \mathbb{R}$,

1. Each agent $i$ *declares* a valuation $v'_i : A \to \mathbb{R}$,

2. A **social choice function** $f : V_1 \times \ldots V_n \to A$ is applied on $(v'_i)_i$. 
An **outcome** is to be chosen by a *central authority* from a set $A$. There are $n$ agents with different preferences.

0. Each agent $i$ has a true **valuation** $v_i : A \to \mathbb{R}$,
1. Each agent $i$ **declares** a valuation $v'_i : A \to \mathbb{R}$,
2. A **social choice function** $f : V_1 \times \ldots \times V_n \to A$ is applied on $(v'_i)_i$.

**Currency**

To prevent manipulations, we need a common currency in which agents can **pay taxes** or be **paid**.

Given outcome $a$, if Player $i$ pays some quantity $m$ of money, then his **utility** is $v_i(a) - m$. 
Direct Revelation Mechanisms

An **outcome** is to be chosen by a *central authority* from a set $A$. There are $n$ agents with different preferences.

0. Each agent $i$ has a true **valuation** $v_i : A \to \mathbb{R}$,
1. Each agent $i$ declares a valuation $v'_i : A \to \mathbb{R}$,
2. A **social choice function** $f : V_1 \times \ldots V_n \to A$ is applied on $(v'_i)_i$.
3. Each agent $i$ pays an amount determined by a **payment function** $p_i : V_1 \times \ldots V_n \to \mathbb{R}$.

**Currency**

To prevent manipulations, we need a common currency in which agents can **pay taxes** or be **paid**.

Given outcome $a$, if Player $i$ pays some quantity $m$ of money, then his **utility** is $v_i(a) - m$.

**Mechanism**: $(f, p_1, \ldots, p_n)$. 
Example

Sell an item to the highest bidder.

**Objective:** No one should increase his utility by lying.

**Regular Auction**

An item is to be sold to the highest bidder. \( A = \text{set of bidders.} \)

- Each player has a private value \( v_i \), and declares a value \( v'_i \).
- Payment functions: The highest bidder \( i_0 \) pays \( v'_{i_0} \), others pay 0.
- Utilities: For the highest bidder \( i_0 \) the utility is \( v_{i_0} - v'_{i_0} \), for others it is 0.
Example

Sell an item to the highest bidder.

**Objective:** No one should increase his utility by lying.

### Regular Auction

An item is to be sold to the highest bidder. \( A = \text{set of bidders.} \)

- Each player has a private value \( v_i \), and declares a value \( v'_i \).
- Payment functions: The highest bidder \( i_0 \) pays \( v'_i \), others pay 0.
- Utilities: For the highest bidder \( i_0 \) the utility is \( v_i - v'_i \), for others it is 0.

Assume every one is telling the truth, and \( v_1 < v_2 < \ldots < v_{n-1} < v_n \).

Outcome: \( n \) gets the item and pays \( v_n \). Utility: 0.
But he would rather lie and declare \( v_{n-1} + \epsilon \), and have utility \( v_n - v_{n-1} - \epsilon \).
Example

Sell an item to the highest bidder.

**Objective:** No one should increase his utility by lying.

**Regular Auction**

An item is to be sold to the highest bidder. $A = \text{set of bidders}.$

- Each player has a private value $v_i,$ and declares a value $v'_i.$
- Payment functions: The highest bidder $i_0$ pays $v'_{i_0},$ others pay 0.
- Utilities: For the highest bidder $i_0$ the utility is $v_{i_0} - v'_{i_0},$ for others it is 0.

Assume every one is telling the truth, and $v_1 < v_2 < \ldots < v_{n-1} < v_n.$

Outcome: $n$ gets the item and pays $v_n.$ Utility: 0.
But he would rather lie and declare $v_{n-1} + \epsilon,$ and have utility $v_n - v_{n-1} - \epsilon.$

Open to manipulation! Players will try to learn others’ valuations.
Objective & Constraints

Social Welfare

We focus on social choice functions that maximize social welfare, i.e.

\[ f(\bar{v}) \in \text{argmax}_{a \in A} \sum_{i=1}^{n} v_i(a). \]

Truthfulness

Given \( f \), we want to find \( p_1, \ldots, p_n \) such that

\[ v_i(f(\bar{v})) - p_i(\bar{v}) \geq v_i(f(\bar{v}_{-i}, v_i')) - p_i(\bar{v}_{-i}, v_i'). \]

Telling the truth should be a dominant strategy (may not be unique).

Such a function is implementable: agents do not have incentive to cheat.
Without payments, what social choice functions are truthful?

**Gibbard-Satterthwaite**

Let $f$ be a truthful social choice function onto $A$, with $|A| \geq 3$, then $f$ is a dictatorship. (the choice is dictated by one of the players).

Money and payments, or other assumptions are necessary to truthfully implement non-trivial functions.
Vickrey-Clarke-Groves (VCG) Mechanisms

VCG payments

Consider a social choice function $f$ that maximizes social welfare. Define payments by

- $p_i(\vec{v}) = h_i(\vec{v}_{-i}) - \sum_{j \neq i} v_j(f(\vec{v}))$,
- where $h_i$ are arbitrary functions (not depending on $v_i$).

Intuition:
- Each player is paid the social welfare of the other players.
  $\Rightarrow$ has incentive to increase others’ welfare.
- A player’s own declaration does not affect $h_i$. The payment is reduced only by increasing others’ welfare.
Theorem

Consider a mechanism \((f, p_1, \ldots, p_n)\) which maximizes social welfare, and payments are given by VCG (for any choice of \(h_i\)). The mechanism is truthful.

Proof

For any player \(i\) with valuation \(v_i\), and declarations \(\tilde{v}_{-i}\) of others, prove that declaring \(v_i\) is as good as declaring any other \(v_i'\).

Denote \(a = f(v_i, \tilde{v}_{-i})\) and \(a' = f(v_i', \tilde{v}_{-i})\).
Theorem

Consider a mechanism \((f, p_1, \ldots, p_n)\) which maximizes social welfare, and payments are given by VCG (for any choice of \(h_i\)). The mechanism is truthful.

Proof

For any player \(i\) with valuation \(v_i\), and declarations \(\vec{v}_{-i}\) of others, prove that declaring \(v_i\) is as good as declaring any other \(v_i'\).

Denote \(a = f(v_i, \vec{v}_{-i})\) and \(a' = f(v_i', \vec{v}_{-i})\).

The utility of declaring \(v_i\) is \(v_i(a) + \sum_{j \neq i} v_j(a) - h_i(\vec{v}_{-i})\).

Declaring \(v_i'\) yields \(v_i(a') + \sum_{j \neq i} v_j(a') - h_i(\vec{v}_{-i})\).
**Theorem**

Consider a mechanism \((f, p_1, \ldots, p_n)\) which maximizes social welfare, and payments are given by VCG (for any choice of \(h_i\)). The mechanism is truthful.

**Proof**

For any player \(i\) with valuation \(v_i\), and declarations \(\vec{v}_{-i}\) of others, prove that declaring \(v_i\) is as good as declaring any other \(v'_i\).

Denote \(a = f(v_i, \vec{v}_{-i})\) and \(a' = f(v'_i, \vec{v}_{-i})\).

The utility of declaring \(v_i\) is \(v_i(a) + \sum_{j \neq i} v_j(a) - h_i(\vec{v}_{-i})\).

Declaring \(v'_i\) yields \(v_i(a') + \sum_{j \neq i} v_j(a') - h_i(\vec{v}_{-i})\).

But \(a = f(v_i, \vec{v}_{-i})\) maximizes social welfare, so

\[
    v_i(a) + \sum_{j \neq i} v_j(a) - h_i(\vec{v}_{-i}) \geq v_i(a') + \sum_{j \neq i} v_j(a') - h_i(\vec{v}_{-i}).
\]
Vickrey Auction

An item is to be sold to the highest bidder.

- Each player has a private value $v_i$, and declares a value $w_i$.
- Payment functions: The highest bidder $i_0$ pays $\max_{i \neq i_0} w_i$, others pay 0.
- Utilities: For the highest bidder $i_0$ the utility is $v_{i_0} - w_{i_0}$, for others it is 0.
Vickrey Auction

An item is to be sold to the highest bidder.

- Each player has a private value $v_i$, and declares a value $w_i$.
- Payment functions: The highest bidder $i_0$ pays $\max_{i \neq i_0} w_i$, others pay 0.
- Utilities: For the highest bidder $i_0$ the utility is $v_{i_0} - w_{i_0}$, for others it is 0.

Assume every one is telling the truth, and $v_1 < v_2 < \ldots < v_{n-1} < v_n$.

Outcome: $n$ gets the item and pays $v_{n-1}$. Utility: $v_n - v_{n-1}$.

Would lying help?

- If $n$ declares $v'_n > v_n$, he would still get the item and pay $v_{n-1}$. $\Rightarrow$ same utility.
- If $n$ declares $v'_n \in (v_{n-1}, v_n)$, same utility.
- If $n$ declares $v'_n \leq v_{n-1}$, he won’t get the item. $\Rightarrow$ utility 0.
Vickrey Auction

An item is to be sold to the highest bidder.

- Each player has a private value $v_i$, and declares a value $w_i$.
- Payment functions: The highest bidder $i_0$ pays $\max_{i \neq i_0} w_i$, others pay 0.
- Utilities: For the highest bidder $i_0$ the utility is $v_{i_0} - w_{i_0}$, for others it is 0.

In fact, payments in Vickrey’s auction are given by a VCG mechanism.

- Giving the item to the highest bidder maximizes social welfare i.e. $\sum_i v_i(a)$.
- Highest bidder $n$ pays

$$p_n(\vec{v}) = h_i(\vec{v}_{-i}) - \sum_{j \neq i} v_j(f(\vec{v})) = v_{n-1} + 0$$

Others pay 0.
The choice of $h_i$ is rather arbitrary.

- A mechanism is **individually rational** if always $v_i(f(\vec{v})) - p_i(\vec{v}) \geq 0$.
- A mechanism is **feasible** if always $\sum_i p_i(\vec{v}) \geq 0$ (mechanism does not need external financing).

**Clarke’s Pivot Rule:** Choose

$$p_i(\vec{v}) = \max_b \sum_{j \neq i} v_j(b) - \sum_{j \neq i} v_j(a),$$

where $a = f(\vec{v})$.

Each player $i$ pays the damage they cause to others.

Clarke’s Pivot Rule $\Rightarrow$ all VCG mechanisms are ind. rational and feasible.
**Goal:** Allocate the shortest (cheapest) path from $s$ to $t$.
- Edges are owned by different selfish agents.
- Each edge $e$ costs $c_e$ to the owner if it is taken, 0 otherwise.

Maximizing social welfare: finding the shortest path ($\sum_{e \in p} c_e$).

**VCG:** To each $e_0 \in p$, pay $\sum_{e \in p'} c_e - \sum_{e \in p \setminus \{e_0\}} c_e$, where $p$ is the shortest path, and $p'$ the shortest path not using $e_0$. 
Affine maximizer

A social choice function $f$ is an **affine maximizer** if for some $A' \subset A$, and weights $w_1, \ldots, w_n > 0$, and $c_a \in \mathbb{R}$ for $a \in A'$, we have

$$f(v_1, \ldots, v_n) \in \arg\max_{a \in A'} (c_a + \sum_{i} w_i v_i(a)),$$

Payment functions can be adapted naturally with the weights.

**Theorem [Roberts 1979]**

If $|A| \geq 3$ and $f$ is onto, $V_i = \mathbb{R}^A$, and $(f, p_1, \ldots, p_n)$ is truthful, then $f$ is an affine maximizer.

Not true for $|A| \leq 2$.

**Open question:** Relaxing the hypothesis $V_i = \mathbb{R}^A$. 
Why Payments?

Reminder:

**Gibbard-Satterthwaite**

Let $f : V_1 \times \ldots \times V_n \rightarrow A$ be a truthful social choice function onto $A$, with $|A| \geq 3$, and $V_i = \mathbb{R}^A$, then $f$ is a dictatorship.

For $|A| = 2$?

Affine maximizers are not the only truthful functions if $V_i \neq \mathbb{R}^A$.

In the rest: restrictions of $V_i$. 
Direct Characterization of Truthfulness

(Relevant when $|A| \leq 2$ or $V_i \subseteq \mathbb{R}^A$).

A mechanism $(f, p_1, \ldots, p_n)$ is truthful iff the following holds for each $i$ and $\vec{v}$,

1. The payment $p_i$ only depends on $f(\vec{v})$ and $\vec{v}_{-i}$ (and not on $v_i$ in any other way).
   If we fix $\vec{v}_{-i}$, the payment $p_a$ only depends on the outcome $a$.

2. The mechanism optimizes for each player:

   $$ f(v_i, \vec{v}_{-i}) \in \arg \max_a (v_i(a) - p_a) $$

   where $a$ ranges over $f(\cdot, \vec{v}_{-i})$.

Proof.
In computer science, objectives different than maximizing social welfare arise, e.g. truthful scheduling.

Several tasks need to be scheduled on different machines.
- Each machine declares its available processing time,
- and is a selfish agent, trying to avoid work.

We want to schedule with minimum makespan. If \( A \) is the set of schedulings for machines \( M \):

\[
\min_{a \in A} \max_{i \in M} \text{load}(i).
\]
A Scheduling on Related Machines

There $n$ jobs to be assigned to $m$ machines. Job $j$ consumes $p_j$ time units, and machine $i$ has speed $c_i$. (So machine $i$ requires $p_j c_i$ time to complete job $j$).

The load of machine $i$, is

$$l_i = \sum_{j|j \text{ assigned to } i} p_j.$$

Under payments $P_i$, the utility of machine $i$ is

$$-l_i c_i - P_i.$$

Machines are agents; they declare $c_i$.

**Objective:** Design a truthful mechanism selecting a scheduling with makespan $\min_a \max_i l^a_i c_i$

Valuations are “one-dimensional” ($c_i$ determines the valuation).
Characterization of Truthfulness: Weak Monotonicity

**Weak monotonicity**

A social choice function satisfies **weak monotonicity** (WMON) iff for all $i$ and $\vec{v}_-i$,

$$f(v_i, \vec{v}_-i) = a \neq b = f(v'_i, \vec{v}_-i)$$

implies

$$v_i(a) - v_i(b) \geq v'_i(a) - v'_i(b).$$

Recall that mechanism $(f, p_1, \ldots, p_n)$ is truthful iff

$$v_i(f(\vec{v})) - p_i(\vec{v}) \geq v_i(f(\vec{v}_-i, v'_i)) - p_i(\vec{v}_-i, v'_i).$$

**Theorem**

- If $(f, p_1, \ldots, p_n)$ is truthful, then $f$ satisfies WMON.
- If $f$ satisfies WMON, there exist $p_1, \ldots, p_n$ such that the mechanism is truthful. (Holds whenever valuation sets are convex sets).

**Proof.**
Application of WMON

Fix $\bar{c}_{-i}$, costs declared by all players but $i$. Then, $l_i(c_i)$ is a function of $c_i$.

Consider $c_i < c_i'$. WMON means

\[-l_i(c_i)c_i + l_i(c_i')c_i \geq -l_i(c_i)c'_i + l_i(c'_i)c'_i \iff l_i(c'_i) \leq l_i(c_i).\]

which means that the work load should be nonincreasing!
WMON implies implementability (Proof by figure)

Assume an algorithm with nonincreasing load curves is given. Consider payments $p_i(c) = \int_0^c l_i(x)dx - c \cdot l_i(c)$.

Utility (with declared cost $c'$):

$$-c \cdot l_i(c') - p_i(c) = -\int_0^c l_i(x)dx$$
WMON implies implementability (Proof by figure)

Assume an algorithm with nonincreasing load curves is given. Consider payments $p_i(c) = \int_0^c l_i(x)dx - c \cdot l_i(c)$.

Utility (with declared and true cost $c$): $- c \cdot l_i(c) - p_i(c) = - \int_0^c l_i(x)dx$

Utility (with declared cost $c'$): $- c \cdot l_i(c') - (p_i(c) + A)$
To prevent utilities to be negative, consider.

**Theorem [Archer, Tardos FOCS’01]**

A scheduling algorithm is truthfully implementable iff its load functions are nonincreasing.

In this case, the following payments functions yield individually rational dominant strategy implementations:

\[
p_i(c) = \int_0^c [l_i(x) - l_i(c)] \, dx + \int_c^{\infty} l_i(x) \, dx.
\]
Overall Theorem

To prevent utilities to be negative, consider.

Theorem [Archer, Tardos FOCS’01]

A scheduling algorithm is truthfully implementable iff its load functions are nonincreasing.

In this case, the following payments functions yield individually rational dominant strategy implementations:

\[ p_i(c) = \int_0^c [l_i(x) - l_i(c)] \, dx + \int_c^\infty l_i(x) \, dx. \]

How to design an efficient algorithm with nonincreasing load functions? (Efficient computation of a scheduling and payments)
Existing Algorithms and The New One

The problem is NP-complete [Hochbaum & Shmoys 88]. We want everything to be computable in PTIME.

A PTAS is known (for all \( \epsilon > 0 \), an \( \epsilon \)-approximation algorithm) but is not nonincreasing!

**Truthful Algorithm (sketch):**

- Fix a time bound \( T \)
- Create a bin of size \( T/c_i \) for for machine \( i \)
- Greedily solve the bin packing problem: assign longest job to fastest machine upto time \( T \), and cut fractionally
- Randomly round fractional jobs

For an efficiently computable \( T \), this gives a factor-2 approximation alg. that is truthful in expectation.
Let us fix $A = [0, 1]$. A relation $\preceq$ on $A$ is \textbf{single-peaked} if there exists $p \in A$ such that

$$\forall x \in A \setminus \{p\}, x \prec \lambda x + (1 - \lambda)p.$$
Let us fix $A = [0, 1]$. A relation $\preceq$ on $A$ is **single-peaked** if there exists $p \in A$ such that

$$\forall x \in A \setminus \{p\}, x \prec \lambda x + (1 - \lambda)p.$$ 

### Examples

1. **Average of peaks:**

   $$f(\preceq_1, \ldots, \preceq_n) = \frac{1}{n} \sum_i p_i.$$
Let us fix $A = [0, 1]$. 
A relation $\preceq$ on $A$ is **single-peaked** if there exists $p \in A$ such that 
\[ \forall x \in A \setminus \{p\}, \ x \prec \lambda x + (1 - \lambda)p. \]

**Examples**

1. Average of peaks:
   \[ f(\preceq_1, \ldots, \preceq_n) = \frac{1}{n} \sum_i p_i. \]

2. Dictatorship:
   \[ f(\preceq_1, \ldots, \preceq_n) = p_{i_0}. \]
Let us fix $A = [0, 1]$. A relation $\preceq$ on $A$ is **single-peaked** if there exists $p \in A$ such that

$$\forall x \in A \setminus \{p\}, x \prec \lambda x + (1 - \lambda)p.$$ 

**Examples**

1. **Average of peaks:**

   $$f(\preceq_1, \ldots, \preceq_n) = \frac{1}{n} \sum_i p_i.$$ 

2. **Dictatorship:**

   $$f(\preceq_1, \ldots, \preceq_n) = p_{i_0}.$$ 

3. **Median of peaks:**

   $$f(\preceq_1, \ldots, \preceq_n) = \text{median}\{p_i\}_i.$$
Under single-peaked preferences, a social choice function $f$ is onto, anonymous, and truthful if and only if $\exists y_1, \ldots, y_{n-1} \in [0, 1]$ such that

$$f(\preceq_1, \ldots, \preceq_n) = \text{median}(p_1, \ldots, p_n, y_1, \ldots, y_{n-1}).$$

anonymous: invariant under permutation of indices.

A characterization is available as “generalized median voter schemes” for onto and truthful functions.

Truthfulness is not the only solution concept. One can design games with e.g. Nash equilibria.

And apply it on richer game structures

Combinatorial auctions! (e.g. Google Ads)


Archer, Tardos. Truthful Mechanisms for one-parameter Agents. FOCS’01


Nisan et al. Algorithmic Mechanism Design. (LSV library).