Revelation Principle; Quasilinear Utility

Lecture 14

Lecture Overview

1. Recap
2. Revelation Principle
3. Impossibility
4. Quasilinear Utility
5. Risk Attitudes

Bayesian Game Setting

- Extend the social choice setting to a new setting where agents can't be relied upon to disclose their preferences honestly.
- Start with a set of agents in a Bayesian game setting (but no actions).

Definition (Bayesian game setting)

A Bayesian game setting is a tuple (N, O, Θ, p, u), where

- N is a finite set of n agents;
- O is a set of outcomes;
- $\Theta = \Theta_1 \times \cdots \times \Theta_n$ is a set of possible joint type vectors;
- p is a (common prior) probability distribution on Θ; and
- $u = (u_1, \ldots, u_n)$, where $u_i : O \times \Theta \mapsto \mathbb{R}$ is the utility function for each player i.

Mechanism Design

Definition (Mechanism)

A mechanism (for a Bayesian game setting (N, O, Θ, p, u)) is a pair (A, M), where

- $A = A_1 \times \cdots \times A_n$, where A_i is the set of actions available to agent $i \in N$; and
- $M : A \mapsto \Pi(O)$ maps each action profile to a distribution over outcomes.

Thus, the designer gets to specify

- the action sets for the agents (though they may be constrained by the environment)
- the mapping to outcomes, over which agents have utility
- can't change outcomes; agents' preferences or type spaces

Implementation in Dominant Strategies

Definition (Implementation in dominant strategies)

Given a Bayesian game setting (N, O, Θ, p, u), a mechanism (A, M) is an implementation in dominant strategies of a social choice function C (over N and O) if for any vector of utility functions u, the game has an equilibrium in dominant strategies, and in any such equilibrium a^* we have $M(a^*) = C(u)$.

Implementation in Bayes-Nash equilibrium

Definition (Bayes–Nash implementation)

Given a Bayesian game setting (N, O, Θ, p, u), a mechanism (A, M) is an implementation in Bayes–Nash equilibrium of a social choice function C (over N and O) if there exists a Bayes–Nash equilibrium of the game of incomplete information (N, A, Θ, p, u) such that for every $\theta \in \Theta$ and every action profile $a \in A$ that can arise given type profile θ in this equilibrium, we have that $M(a) = C(u(\cdot, \theta))$.

Properties

Forms of implementation

- Direct Implementation: agents each simultaneously send a single message to the center
- Indirect Implementation: agents may send a sequence of messages; in between, information may be (partially) revealed about the messages that were sent previously like extensive form

We can also insist that our mechanism satisfy properties like the following:

- individual rationality: agents are better off playing than not playing
- budget balance: the mechanism gives away and collects the same amounts of money
- truthfulness: agents honestly report their types

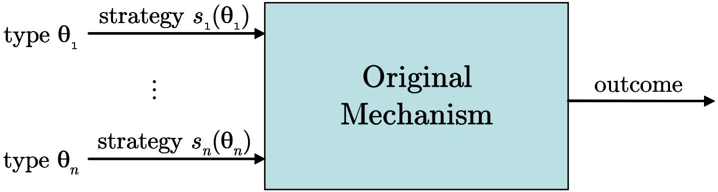
Lecture Overview

1. Recap
2. Revelation Principle
3. Impossibility
4. Quasilinear Utility
5. Risk Attitudes

Revelation Principle

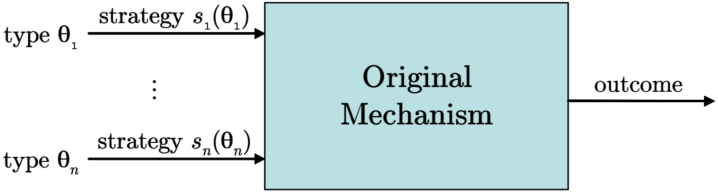
- It turns out that truthfulness can always be achieved!
- Consider an arbitrary, non-truthful mechanism (e.g., may be indirect)

Revelation Principle



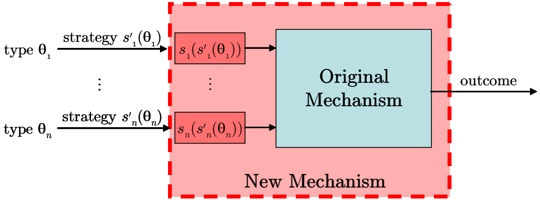
- It turns out that truthfulness can always be achieved!
- Consider an arbitrary, non-truthful mechanism (e.g., may be indirect)

Revelation Principle



- It turns out that truthfulness can always be achieved!
- Consider an arbitrary, non-truthful mechanism (e.g., may be indirect)
- Recall that a mechanism defines a game, and consider an equilibrium $s = (s_1, \ldots, s_n)$

Revelation Principle



- We can construct a new direct mechanism, as shown above
- This mechanism is truthful by exactly the same argument that s was an equilibrium in the original mechanism
- "The agents don't have to lie, because the mechanism already lies for them."

Computational Criticism of the Revelation Principle

- computation is pushed onto the center
  - often, agents' strategies will be computationally expensive
    - e.g., in the shortest path problem, agents may need to compute shortest paths, cutsets in the graph, etc.
  - since the center plays equilibrium strategies for the agents, the center now incurs this cost
- if computation is intractable, so that it cannot be performed by agents, then in a sense the revelation principle doesn't hold
  - agents can't play the equilibrium strategy in the original mechanism
  - however, in this case it's unclear what the agents will do

Discussion of the Revelation Principle

- The set of equilibria is not always the same in the original mechanism and revelation mechanism
  - of course, we've shown that the revelation mechanism does have the original equilibrium of interest
  - however, in the case of indirect mechanisms, even if the indirect mechanism had a unique equilibrium, the revelation mechanism can also have new, bad equilibria
- So what is the revelation principle good for?
  - recognition that truthfulness is not a restrictive assumption
  - for analysis purposes, we can consider only truthful mechanisms, and be assured that such a mechanism exists
  - recognition that indirect mechanisms can't do (inherently) better than direct mechanisms

Lecture Overview

Impossibility Result

Theorem (Gibbard-Satterthwaite)

Consider any social choice function C of N and O. If:

1. $|O| \geq 3$ *(there are at least three outcomes);*
2. C *is onto; that is, for every* $o \in O$ *there is a preference profile* $[\succ]$ *such that* $C([\succ]) = o$ *(this property is sometimes also called citizen sovereignty); and*
3. C *is dominant-strategy truthful,*

then C is dictatorial.

What does this mean?

- We should be discouraged about the possibility of implementing arbitrary social-choice functions in mechanisms.
- However, in practice we can circumvent the Gibbard-Satterthwaite theorem in two ways:
  - use a weaker form of implementation
    - note: the result only holds for dominant strategy implementation, not e.g., Bayes-Nash implementation
  - relax the onto condition and the (implicit) assumption that agents are allowed to hold arbitrary preferences

Lecture Overview

Quasilinear Utility

Definition (Quasilinear preferences)

Agents have quasilinear preferences in an n-player Bayesian game when the set of outcomes is

$$O = X \times \mathbb{R}^n$$

for a finite set X, and the utility of an agent i given joint type θ is given by

$$u_i(o, \theta) = u_i(x, \theta) - f_i(p_i),$$

where $o = (x, p)$ is an element of O, $u_i : X \times \Theta \mapsto \mathbb{R}$ is an arbitrary function and $f_i : \mathbb{R} \mapsto \mathbb{R}$ is a strictly monotonically increasing function.

Quasilinear utility

- $u_i(o, \theta) = u_i(x, \theta) - f_i(p_i)$
- We split the mechanism into a choice rule and a payment rule:
  - $x \in X$ is a discrete, non-monetary outcome
  - $p_i \in \mathbb{R}$ is a monetary payment (possibly negative) that agent i must make to the mechanism
- Implications:

Quasilinear utility

- $u_i(o, \theta) = u_i(x, \theta) - f_i(p_i)$
- We split the mechanism into a choice rule and a payment rule:
  - $x \in X$ is a discrete, non-monetary outcome
  - $p_i \in \mathbb{R}$ is a monetary payment (possibly negative) that agent i must make to the mechanism
- Implications:
  - $u_i(x, \theta)$ is not influenced by the amount of money an agent has
  - agents don't care how much others are made to pay (though they *can* care about how the choice affects others.)

Quasilinear utility

- $u_i(o,\theta) = u_i(x,\theta) - f_i(p_i)$
- We split the mechanism into a choice rule and a payment rule:
  - $x \in X$ is a discrete, non-monetary outcome
  - $p_i \in \mathbb{R}$ is a monetary payment (possibly negative) that agent i must make to the mechanism
- Implications:
  - $u_i(x,\theta)$ is not influenced by the amount of money an agent has
  - agents don't care how much others are made to pay (though they *can* care about how the choice affects others.)
- What is $f_i(p_i)$?

Lecture Overview

Fun game

- Look at your piece of paper: it contains an integer x
- Go around the room offering everyone the following gamble:
    - they pay you x
    - you flip a coin:
        - heads: they win and get paid $2x$
        - tails: they lose and get nothing.
    - Players can accept the gamble or decline.
        - Answer honestly (imagining the amounts of money are real)
        - play the gamble to see what would have happened.
    - Keep track of:
        - Your own "bank balance" from others' gambles you accepted.
        - The number of people who accepted your offer.

Risk Attitudes

- How much is \$1 worth?
  - What are the units in which this question should be answered?

Risk Attitudes

- How much is $1 worth?
  - What are the units in which this question should be answered? Utils (units of utility)

Risk Attitudes

- How much is \$1 worth?
  - What are the units in which this question should be answered? Utils (units of utility)
  - Different amounts depending on the amount of money you already have

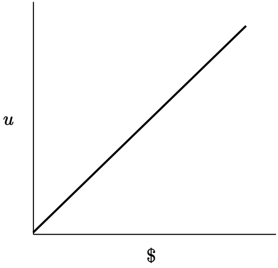
Risk Attitudes

- How much is $1 worth?
  - What are the units in which this question should be answered? Utils (units of utility)
  - Different amounts depending on the amount of money you already have
- How much is a gamble with an expected value of $1 worth?

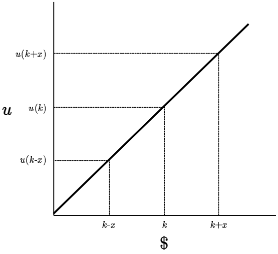
Risk Attitudes

- How much is $1 worth?
  - What are the units in which this question should be answered? Utils (units of utility)
  - Different amounts depending on the amount of money you already have
- How much is a gamble with an expected value of $1 worth?
  - Possibly different amounts, depending on how risky it is

Risk Neutrality

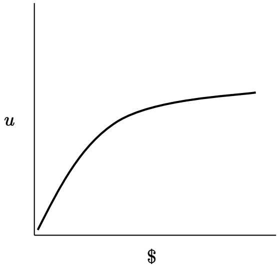


(a) Risk neutrality

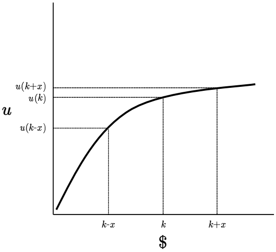


(b) Risk neutrality: fair lottery

Risk Aversion

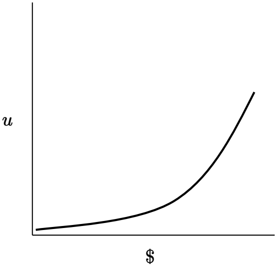


(c) Risk aversion

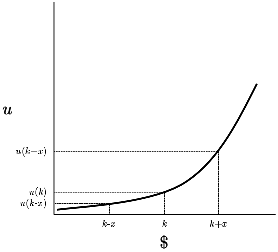


(d) Risk aversion: fair lottery

Risk Seeking



(e) Risk seeking



(f) Risk seeking: fair lottery