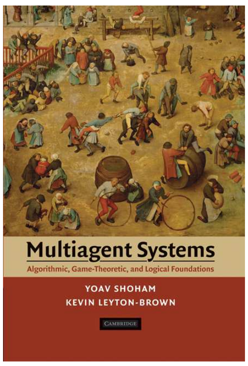
Intelligent Autonomous Agents and Cognitive Robotics

Topic 12: Mechanism Design

Ralf Möller, Rainer Marrone
Hamburg University of Technology

Acknowledgement

Material from CS 886
Advanced Topics in AI Electronic Market Design
Kate Larson
Waterloo Univ.

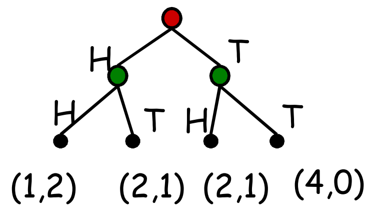


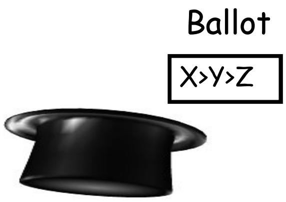
2

Introduction

So far we have looked at

- Game Theory
  - Given a game we are able to analyze the strategies agents will follow
- Social Choice Theory
  - Given a set of agents' preferences we can choose some outcome





3

Introduction

- Now: Mechanism Design
  - Game Theory + Social Choice
- Goal of Mechanism Design is to
  - Obtain a dedicated outcome (function of agents' preferences)
  - But agents are rational
    They may lie about their preferences
- Goal:
  Define the rules of a game so that in equilibrium the agents do what the social community in general wants

4

Fundamentals

- Set of possible outcomes, O
- Agents $i \in I$, $|I|=n$, each agent i has type $\theta i \in \Theta i$
  - Type captures all private information that is relevant to agent's decision making (its payoffs, which may be different)
- Utility $ui(o, \theta i)$, over outcome $o \in O$
- Recall: goal is to implement some **system-wide** solution
  - Captured by a social choice function (SCF)

$$f: \Theta_1 \times \ldots \times \Theta_n \rightarrow O$$

$f(\theta_1,\ldots\theta_n)=o$ is a collective choice

5

Mechanisms

- Recall: We want to implement a social choice function
  - Need to know agents' preferences
  - They may not reveal them to us truthfully
- Example:
  - 1 item to allocate, and want to give it to the agent who values it the most
  - If we just ask agents to tell us their preferences, they may lie

7

Mechanism Design Problem

- By having agents interact through an institution (M) we might be able to solve the problem
- Mechanism:

$$M=(S_1,\dots,S_n,\ g(.))$$

Strategy spaces of agents

Outcome function

$$g:S_1 x\dots x\ S_n \rightarrow O$$

8

Implementation

- A mechanism $M=(S_1,\dots,S_n,g(.))$

implements social choice function $f(\theta)$

if there is an equilibrium strategy profile

$$s^*(.)=(s^*_1(.),\dots,s^*_n(.))$$

of the game induced by M such that

$$g(s_1^*(\theta_1),\dots,s_n^*(\theta_n))=f(\theta_1,\dots,\theta_n)$$

for all

$$(\theta_1,\dots,\theta_n) \in \Theta_1 x\dots\ x\ \Theta_n$$

9

Implementation

- We did not specify the type of equilibrium in the definition
  - (Mixed) Nash
  - Bayes-Nash
  - Dominant

10

Direct Mechanisms

- Recall that a mechanism specifies the strategy sets of the agents
  - These sets can contain complex strategies
- **Direct mechanisms:**
  - Mechanism in which $S_i=\Theta_i$ for all i, and $g(\theta)=f(\theta)$ for all $\theta \in \mathbf{\Theta}_1\mathbf{x}\ldots\mathbf{x}\Theta_n$
- **Incentive-compatible:**
  - A direct mechanism is incentive-compatible if it has an equilibrium s^* where $s^*_i(\theta_i)=\theta_i$ for all $\theta_i \in \mathbf{\Theta}_i$ and all i
  - (truth telling by all agents is an equilibrium)
  - Strategy-proof if dominant-strategy equilibrium

11

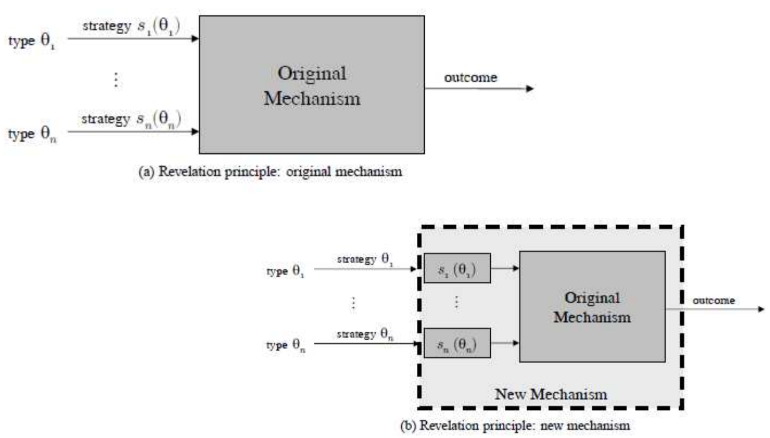
Dominant Strategy Implementation

- Is a certain social choice function implementable in dominant strategies? Did the mechanism enforce dominant strategies?
  - In principle we would need to consider all possible mechanisms

- **Revelation Principle** (for Dom Strategies)
  - Suppose there exists a **(in)direct** mechanism $M=(S_1,\ldots,S_n,g(.))$ that implements social choice function f() in dominant strategies. Then there is a direct strategy-proof mechanism, M', which also implements f().

12

Revelation Principle: Intuition



(a) Revelation principle: original mechanism

(b) Revelation principle: new mechanism

13

Theoretical Implications

- Literal interpretation: Need only study direct mechanisms
    - This is a much smaller space of mechanisms
  - Negative results: If no direct mechanism can implement SCF f() then no mechanism can do it => impossibility theorems, e.g. Arrow in voting.
  - Analysis tool:
    - Best direct mechanism gives us an upper bound on what we can achieve with an indirect mechanism
    - Analyze all direct mechanisms and choose the best one

15

Practical Implications

- Incentive-compatibility is “free” from an implementation perspective
- **BUT!!!**
  - A lot of mechanisms used in practice are not direct and incentive-compatible
  - Maybe there are some issues that are being ignored here

16

Quasi-Linear Preferences

- Outcome $o=(x,t_1,\dots,t_n)$
  - x is a "project choice" and $t_i \in \mathbf{R}$ are transfers (money)
- Utility function of agent i
  - $u_i(o,\theta_i)=u_i((x,t_1,\dots,t_n),\theta_i)=v_i(x,\theta_i)-t_i$
- Quasi-linear mechanism: $M=(S_1,\dots,S_n,g(.))$ where $g(.)=(x(.),t_1(.),\dots,t_n(.))$

20

Social choice functions and quasi-linear settings

- SCF is efficient if for all types $\theta=(\theta_1,\dots,\theta_n)$
  - $\sum^n_{i=1} v_i(x(\theta),\theta_i) \geq \sum^n_{i=1} v_i(x'(\theta),\theta_i) \quad \forall\, x'(\theta)$
  - Aka social welfare maximizing, x is the selection function
- SCF is budget-balanced (BB) if
  - $\sum^n_{i=1} t_i(\theta)=0$
  - Weakly budget-balanced if $\sum^n_{i=1} t_i(\theta)\geq 0$

21

Groves Mechanisms

[Groves 1973]

- A **Groves mechanism**, $M=(S_1,\ldots,S_n, (x,t_1,\ldots,t_n))$ is defined by
  - Choice rule $x^*(\theta')=\text{argmax}_x \sum_i v_i(x,\theta_i')$
  - Transfer rules
    - $t_i(\theta')=h_i(\theta_{-i}')-\sum_{j\neq i} v_j(x^*(\theta'),\theta'_j)$

where $h_i(.)$ is an (arbitrary) function that does not depend on the reported type θ_i' of agent i (quasi linear)

22

VCG Mechanism

(aka Clarke tax mechanism aka Pivotal mechanism)

- Def: Implement efficient outcome,

$$x^*=\text{argmax}_x\sum_i v_i(x,\theta_i')$$

Compute transfers

$$t_i(\theta')=\sum_{j\neq i} v_j(x^{-i},\theta'_j) -\sum_{j\neq i}v_j(x^*, \theta_j')$$

Where $x^{-i}=\text{argmax}_x \sum_{j\neq i}v_j(x,\theta_j')$

VCGs are efficient and strategy-proof

Agent's equilibrium utility is:

$u_i(x^*,t_i,\theta_i)=v_i(x^*,\theta_i)-[\sum_{j\neq i} v_j(x^{-i},\theta_j) -\sum_{j\neq i}v_j(x^*,\theta_j)]$

$= \sum_j v_j(x^*,\theta_j) - \sum_{j\neq i} v_j(x^{-i},\theta_j)$

= marginal contribution to the welfare of the system 24

Example: Building a pool

- The cost of building the pool is \$300
- If together all agents value the pool more than \$300 then it will be built
- VCG Mechanism:
  - Each agent announces their value, v_i
  - If $\sum v_i \geq 300$ then it is built
  - Payments $t_i(\theta_i') = \sum_{j \neq i} v_j(x^{-i}, \theta_j') - \sum_{j \neq i} v_j(x^*, \theta_j')$ if built, 0 otherwise

v1=50, v2=50, v3=250

Pool should be built

$t_1=(250+50)-(250+50)=0$
$t_2=(250+50)-(250+50)=0$
$t_3=(0)-(50+50)=-100$

Not budget balanced

25

Example

- The government is deciding on number of street lights to be installed.
- Three beneficiaries - A, B, C.
- Four alternatives: n = 0, 1, 2, 3 where n is the number of street lights. The cost of a street light is 120.
- The government's objective to install the socially efficient number of street lights.

26

Net benefits with equal cost share

- If n = 2, the total cost is 240.
  Hence, cost share for each is 80 (40 for each lamp).

Resident	No. of street lights			
	0	1	2	3
A	0	60	90	155
B	0	80	120	140
C	0	120	200	220
Cost	0	120	240	360

27

Net benefits with equal cost share

- The private net benefit for A is then 90 − 80 = 10.
- Similarly for B and C and n = 1, 3. Figure show the benefits for each agent.

Resident	No. of street lights			
	0	1	2	3
A	0	60	90	155
B	0	80	120	140
C	0	120	200	220
Cost	0	120	240	360

Resident	No. of street lights			
	0	1	2	3
A	0	20	10	35
B	0	40	40	20
C	0	80	120	100
Social net benefit	0	140	170	155

28

Groves Clarke Taxes

- Is Person A pivotal? Does he has to pay a tax?

Resident	No. of street lights			
	0	1	2	3
A	0	20	10	35
B	0	40	40	20
C	0	80	120	100
Social net benefit	0	140	170	155

Resident	No. of street lights			
	0	1	2	3
B	0	40	40	20
C	0	80	120	100
Social net benefit	0	120	160	120

Person A is not pivotal. Without him, the net benefit is maximum at n = 2. With him the net benefit is maximum at n = 2. So his tax is zero.

29

Person B

Resident	No. of street lights			
	0	1	2	3
A	0	20	10	35
B	0	40	40	20
C	0	80	120	100
Social net benefit	0	140	170	155

Resident	No. of street lights			
	0	1	2	3
A	0	20	10	35
C	0	80	120	100
Social net benefit	0	100	130	135

- Person B however is pivotal. With him the net benefit is maximum at n = 2. Without him the net benefit is maximum at n = 3.
- B's tax is the difference between the sum of net benefits of others at n = 3 and the sum of net benefits of others at n = 2, i.e. 135 − 130 = 5.
- B is paying the tax because his report changes the decision from n = 3 to n = 2.

30

Person C

Resident	No. of street lights			
	0	1	2	3
A	0	20	10	35
B	0	40	40	20
C	0	80	120	100
Social net benefit	0	140	170	155

Resident	No. of street lights			
	0	1	2	3
A	0	20	10	35
B	0	40	40	20
Social net benefit	0	60	50	55

- Person C is pivotal as well. With him the net benefit is maximum at n = 2. Without him the net benefit is maximum at n = 1
- C's tax is therefore the sum of others' benefits at n = 1 and the sum of others' benefits at n = 2, i.e. 60 − 50 = 10.

31

Net benefits with taxes

Resident	No. of street lights				Tax
	0	1	2	3	
A	0	20	10	35	0
B	0	40	40	20	5
C	0	80	120	100	10
Social net benefit	0	140	170	155	

- Post tax net benefit from this scheme:
  10 for A,
  40 − 5 = 35 for B,
  120 − 10 = 110 for C.

32

Incentives for truthful revelation

Resident	No. of street lights			
	0	1	2	3
A	0	20	10	35 ← 70
B	0	40	40	20
C	0	80	120	100
Social net benefit	0	140	170	190

?

- Notice that A's net benefit is maximum at n = 3. Does he have an incentive to lie and change the decision to n = 3?
- Suppose A states his net benefit from n = 3 to be 70 instead of 35. Then, sum of stated net benefits is maximum at n = 3.

33

Incentives for truthful revelation

Resident	No. of street lights			
	0	1	2	3
A	0	20	10	70
B	0	40	40	20
C	0	80	120	100
Social net benefit	0	140	170	190

Resident	No. of street lights			
	0	1	2	3
B	0	40	40	20
C	0	80	120	100
Social net benefit	0	120	160	120

- But then A becomes pivotal. Without him the sum of net benefits is maximum at n = 2.
  His report changes the decision from n = 2 to n = 3.
- So he has to pay a tax and his tax will be equal to 160 − 120 = 40.

34

Incentives for truthful revelation

- A's net benefit from lying will be
  (Net benefit from n = 3) − Tax
  = 35 − 40
  = −5
- A's net benefit from truthfully reporting is 10.
- Hence A doesn't have incentive to lie.
- You can repeat the same exercise for B and C to verify that they do not have incentive to lie either.

35

Clarke tax mechanism…

- Pros
  - Social welfare maximizing outcome
  - Truth-telling is a dominant strategy
  - Feasible in that it does not need a benefactor ($\sum_i t_i \leq 0$) (not discussed here)

36

Participation Constraints

- Agents can not be forced to participate in a mechanism
  - It must be in their own best interest
- A mechanism is **individually rational** (IR) if an agent's (expected) utility from participating is (weakly) better than what it could get by not participating

37

Participation Constraints

- Can classify mechanisms based on participation constraints
  - Let $u_i^*(\theta_i)$ be an agent's utility if it does not participate and has type θ_i
  - Ex ante IR: An agent must decide to participate before it knows its own type and other types
    - $E_{\theta\in\Theta}[u_i(f(\theta),\theta_i)] \geq E_{\theta_i \in \Theta_i}[u_i^*(\theta_i)]$
  - Interim IR: An agent decides whether to participate once it knows its own type, but no other agent's type
    - $E_{\theta_{-i} \in \Theta_{-i}}[u_i(f(\theta_i,\theta_{-i}),\theta_i)] \geq u_i^*(\theta_i)$
  - Ex post IR: An agent decides whether to participate after it knows everyone's types (after the mechanism has completed)
    - $u_i(f(\theta),\theta_i) \geq u_i^*(\theta_i)$

38

Quick Review

- Gibbard-Satterthwaite
  - Impossible to get non-dictatorial mechanisms if using dominant strategy implementation and general preferences
- Groves
  - Possible to get dominant strategy implementation with quasi-linear utilities
    - Efficient
- Clarke (or VCG)
  - Possible to get dominant strategy implementation with quasi-linear utilities
    - Efficient, interim IR

39

The End

- Exam: 30.03, 9:00, Audimax I
- Remember comments in exercises
- There will be no questions about "Mechanism Design" in the exam!!!.

40