Mechanism Design II

CS 886: Multiagent Systems

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
- Clarke Mechanism:
  - Each agent announces their value, v_i
  - If $\sum v_i \geq 300$ then it is built and each pays 100
  - Payments $t_i(\theta_i') = \sum_{j \neq i} v_j(x^{-i}, \theta_j') - \sum_{j \neq i} v_j(x^*, \theta_i')$ if built, 0 otherwise

v1=50, v2=50, v3=250

After each pay 100

Pool should be built

$t_1=(250+50)-(250+50)=0$

$t_2=(250+50)-(250+50)=0$

$t_3=(0)-(100)=-100$

Not budget balanced

Clarke tax mechanism…

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

Clarke tax mechanism…

Budget balance not maintained (in pool example, generally $\sum_i t_i < 0$)

- Have to burn the excess money that is collected

- Thrm. [Green & Laffont 1979]. Assume agents have quasi-linear utilities $u_i(x,t)=v_i(x)-t_i$ where $v_i(x)$ are arbitrary functions.
  - There is no social choice function f() that is implementable in dominant strategies and is both efficient and budget balanced.

Implementation in Bayes-Nash equilibrium

Implementation in Bayes-Nash equilibrium

- Goal is to design the rules of the game (aka mechanism) so that in **Bayes-Nash** equilibrium $(s_1, \ldots, s_n)$, the outcome of the game is $f(\theta_1,\ldots,\theta_n)$
- Weaker requirement than dominant strategy implementation
  - An agent's best response strategy may depend on others' strategies
    - Agents may benefit from counterspeculating each others'
      - Preferences, rationality, endowments, capabilities…
  - Can accomplish more than under dominant strategy implementation
    - E.g., budget balance & Pareto efficiency (social welfare maximization) under quasilinear preferences …

Expected externality mechanism
[d'Aspremont & Gerard-Varet 79; Arrow 79]

- Like Groves mechanism, but side payment is computed based on agent's revelation v_i, averaging over possible true types of the others v_{-i} *
- Outcome $(x, t_1, t_2, \ldots, t_n)$
- *Quasilinear* preferences: $u_i(x, t_i) = v_i(x) - t_i$
- *Utilitarian* setting: Social welfare maximizing choice
  - Outcome $x(v_1, v_2, \ldots, v_n) = \max_x \sum_i v_i(x)$
- Others' expected welfare when agent i announces v_i is $\xi(v_i) = \int_{v_{-i}} p(v_{-i}) \sum_{j \neq i} v_j(x(v_i, v_{-i}))$
  - Measures change in expected externality as agent i changes her revelation

* Assume that an agent's type is its value function

Expected externality mechanism
[d'Aspremont & Gerard-Varet 79; Arrow 79]

- **Thrm.** Assume quasilinear preferences and statistically independent valuation functions v_i. A utilitarian social choice function f: v -> (x(v), t(v)) can be implemented in Bayes-Nash equilibrium if $t_i(v_i) = \xi(v_i) + h_i(v_{-i})$ for arbitrary function h
- Unlike in dominant strategy implementation, budget balance is achievable
  - Intuitively, have each agent contribute an equal share of others' payments
  - Formally, set $h_i(v_{-i}) = - [1 / (n-1)] \sum_{j \neq i} \xi(v_j)$
- Does not satisfy participation constraints (aka individual rationality constraints) in general
  - Agent might get higher expected utility by not participating

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** if an agent's (expected) utility from participating is (weakly) better than what it could get by not participating

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
  - $E_{\theta \in \Theta}[u_i(f(\theta), \theta_i)] \geq E_{\theta_i \in \Theta_i}[u_i^*(\theta_i)]$
- 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)
  - $u_i(f(\theta), \theta_i) \geq u_i^*(\theta_i)$

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 strat implementation with quasi-linear utilities
    - Efficient, interim IR
- D'AVGA
  - Possible to get Bayesian-Nash implementation with quasi-linear utilities
    - Efficient, budget balanced, ex ante IR

Other mechanisms

- We know what to do with
  - Voting
  - Auctions
  - Public projects
- Are there any other "markets" that are interesting?

Bilateral Trade

- Heart of any exchange
- 2 agents (one buyer, one seller), quasi-linear utilities
- Each agent knows its own value, but not the other's
- Probability distributions are common knowledge

- Want a mechanism that is
  - Ex post budget balanced
  - Ex post Pareto efficient: exchange to occur if $v_b \geq v_s$
  - (Interim) IR: Higher expected utility from participating than by not participating

Myerson-Satterthwaite Thm

- **Thm**: In the bilateral trading problem, no mechanism can implement an ex-post BB, ex post efficient, and interim IR social choice function (even in Bayes-Nash equilibrium).

Proof

- Seller's valuation is s_L w.p. α and s_H w.p. $(1-\alpha)$
- Buyer's valuation is b_L w.p. β and b_H w.p. $(1-\beta)$. Say $b_H > s_H > b_L > s_L$
- By revelation principle, can focus on truthful direct revelation mechanisms
- $p(b,s)$ = probability that car changes hands given revelations b and s
  - Ex post efficiency requires: $p(b,s) = 0$ if ($b = b_L$ and $s = s_H$), otherwise $p(b,s) = 1$
  - Thus, $E[p|b=b_H] = 1$ and $E[p|b = b_L] = \alpha$
  - $E[p|s = s_H] = 1-\beta$ and $E[p|s = s_L] = 1$
- $m(b,s)$ = expected price buyer pays to seller given revelations b and s
  - Since parties are risk neutral, equivalently $m(b,s)$ = actual price buyer pays to seller
  - Since buyer pays what seller gets paid, this maintains budget balance ex post
  - $E[m|b] = (1-\alpha)\, m(b, s_H) + \alpha\, m(b, s_L)$
  - $E[m|s] = (1-\beta)\, m(b_H, s) + \beta\, m(b_L, s)$

Proof

- Individual rationality (IR) requires
  - $b\, E[p|b] - E[m|b] \geq 0$ for $b = b_L, b_H$
  - $E[m|s] - s\, E[p|s] \geq 0$ for $s = s_L, s_H$
- Bayes-Nash incentive compatibility (IC) requires
  - $b\, E[p|b] - E[m|b] \geq b\, E[p|b'] - E[m|b']$ for all b, b'
  - $E[m|s] - s\, E[m|s] \geq E[m|s'] - s\, E[m|s']$ for all s, s'
- Suppose $\alpha=\beta= ½$, $s_L=0$, $s_H=y$, $b_L=x$, $b_H=x+y$, where $0 < 3x < y$. Now,
- $IR(b_L)$: $½ x - [½ m(b_L,s_H) + ½ m(b_L,s_L)] \geq 0$
- $IR(s_H)$: $[½ m(b_H,s_H) + ½ m(b_L,s_H)] - ½ y \geq 0$
- Summing gives $m(b_H,s_H) - m(b_L,s_L) \geq y-x$
- Also, $IC(s_L)$: $[½ m(b_H,s_L) + ½ m(b_L,s_L)] \geq [½ m(b_H,s_H) + ½ m(b_L,s_H)]$
  - I.e., $m(b_H,s_L) - m(b_L,s_H) \geq m(b_H,s_H) - m(b_L,s_L)$
- $IC(b_H)$: $(x+y) - [½ m(b_H,s_H) + ½ m(b_H,s_L)] \geq ½ (x+y) - [½ m(b_L,s_H) + ½ m(b_L,s_L)]$
  - I.e., $x+y \geq m(b_H,s_H) - m(b_L,s_L) + m(b_H,s_L) - m(b_L,s_H)$
  - So, $x+y \geq 2 [m(b_H,s_H) - m(b_L,s_L)] \geq 2(y-x)$. So, $3x \geq y$, contradiction. QED

Does market design matter?

- You often here "The market will take care of "it", if allowed to."
- Myerson-Satterthwaite shows that under reasonable assumptions, the market will **NOT** take care of efficient allocation

- For example, if we introduced a disinterested 3rd party (auctioneer), we could get an efficient allocation