Impossibility of General, Dominant-Strategy Implementation

Game Theory Course:
Jackson, Leyton-Brown & Shoham
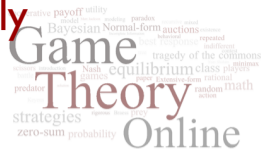
Consider a society N, O and any mechanism A, M for which every agent has a dominant strategy for each preference. There exists a social choice function C (a "direct mechanism") for which truthful announcement of preferences is a dominant strategy.

Consider a society N, O and any mechanism A, M for which every agent has a dominant strategy for each preference. There exists a social choice function C (a "direct mechanism") for which truthful announcement of preferences is a dominant strategy.

So, if we are considering implementation in dominant strategies, it is enough to look only at social choice functions for which truth is a dominant strategy: the set of *non-manipulable* or *strategy-proof* social choice functions.

Impossibility Result

Theorem (Gibbard–Satterthwaite)

Consider a social choice function $C : L^n \mapsto O$. Suppose that

1. *there are at least three outcomes so that $|O| \geq 3$, and*
2. *C is onto; that is, for every $o \in O$ there is a preference profile $[\succ] \in L^n$ such that $C([\succ]) = o$*

Truthful reporting of preferences is a dominant strategy for each agent i and each preference $\succ_i \in L$ if and only if C is dictatorial: there exists i for whom $C([\succ]) = \mathrm{argmax}_O \succ_i$ for all $[\succ] \in L^n$.

Impossibility Result

Theorem (Gibbard–Satterthwaite)

Consider a social choice function $C : L^n \mapsto O$. Suppose that

1. *there are at least three outcomes so that $|O| \geq 3$, and*
2. *C is onto; that is, for every $o \in O$ there is a preference profile $[\succ] \in L^n$ such that $C([\succ]) = o$*

Truthful reporting of preferences is a dominant strategy for each agent i and each preference $\succ_i \in L$ if and only if C is dictatorial: there exists i for whom $C([\succ]) = \text{argmax}_O \succ_i$ for all $[\succ] \in L^n$.

So, any non-dictatorial social choice function on a full domain of preferences and with at least three alternatives will be manipulable by some agents for some preference profiles.

What does this mean?

- Having dominant strategies for all agents and possible preferences is infeasible unless we have a dictatorial social choice function.
- However, in practice we can circumvent the Gibbard–Satterthwaite theorem in various ways:

What does this mean?

- Having dominant strategies for all agents and possible preferences is infeasible unless we have a dictatorial social choice function.
- However, in practice we can circumvent the Gibbard–Satterthwaite theorem in various ways:
  - use a weaker form of implementation:
    - the result only holds for dominant strategy implementation, not e.g., Bayes–Nash implementation

What does this mean?

- Having dominant strategies for all agents and possible preferences is infeasible unless we have a dictatorial social choice function.
- However, in practice we can circumvent the Gibbard–Satterthwaite theorem in various ways:
  - use a weaker form of implementation:
    - the result only holds for dominant strategy implementation, not e.g., Bayes–Nash implementation
  - relax the assumption that agents are allowed to have arbitrary preferences and look at more structured settings.
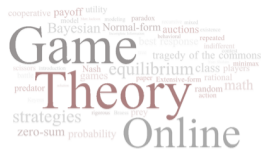
- Single-Peaked domains:
  - median voting
  - or take the max of peaks, or the min of peaks...

Settings with Strategy-Proof Social Choice Functions:

- Single-Peaked domains:
  - median voting
  - or take the max of peaks, or the min of peaks...
- Trade:
  - Have a private value for buying (or selling) an indivisible good
  - A price is fixed in advance,
  - declare whether willing to buy (sell) at that price

Settings with Strategy-Proof Social Choice Functions:

- Single-Peaked domains:
  - median voting
  - or take the max of peaks, or the min of peaks...
- Trade:
  - Have a private value for buying (or selling) an indivisible good
  - A price is fixed in advance,
  - declare whether willing to buy (sell) at that price
- ... we will see more shortly.