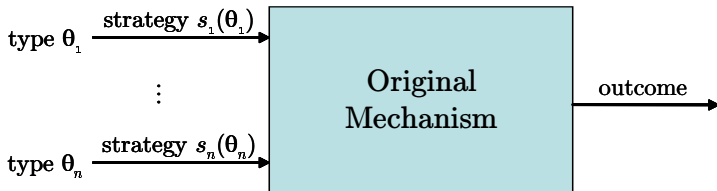


Revelation Principle

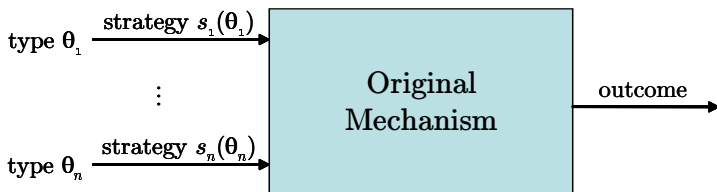
- It turns out that any social choice function that can be implemented by any mechanism can be implemented by a **truthful, direct** mechanism!
- Consider an arbitrary, non-truthful mechanism (e.g., may be indirect)

Revelation Principle



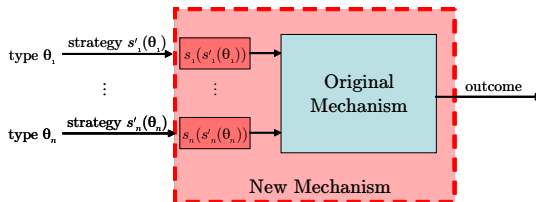
- It turns out that any social choice function that can be implemented by any mechanism can be implemented by a **truthful, direct** mechanism!
- Consider an arbitrary, non-truthful mechanism (e.g., may be indirect)

Revelation Principle



- It turns out that any social choice function that can be implemented by any mechanism can be implemented by a **truthful, direct** mechanism!
- Consider an arbitrary, non-truthful mechanism (e.g., may be indirect)
- Recall that a mechanism defines a game, and consider an equilibrium $s = (s_1, \dots, s_n)$

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

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