CS-698W: Game Theory and Collective Choice **Jul-Nov 2017**

Lecture 16: September 6, 2017

Lecturer: Swaprava Nath *Scribe(s): Sandipan Mandal*

16.1 Recap

In the last class we continued our discussion on **Mechanism Design**. We first discussed *Social Choice Functions (SCF)*, utility functions for SCFs and their types, defined *Mechanism Design (MD)* and its types. Then we went on to discuss *Weakly Dominant Strategy* w.r.t. mechanism and conditions under which a SCF is implemented in dominant startegy. It was followed by discussion on *Strategy Proofs* or *Dominant Strategy Incentive Compatibility (DSIC)* and *Revelation Principle.*
Before going forward lets revisit the definition of DSIC as it will be used later to illustrate a subtle point w.r.t. Bayesian Incentive Compatiblity. The definition is as follows:

Definition 16.1 *A direct mechanism $\langle \Theta, f \rangle$ is* ***Strategy-proof*** *or* ***Dominant Strategy Incentive Compatible (DSIC)*** *if*

$$u_i(f(\theta_i, \tilde{\theta}_{-i}), \theta_i) \geq u_i(f(\theta_i', \tilde{\theta}_{-i}), \theta_i) \qquad \forall \theta_i, \theta_i' \in \Theta_i, \forall \tilde{\theta}_{-i} \in \Theta_{-i}, \forall i \in N \tag{16.1}$$

This definition states that irrespective of whether others are reporting their types to the central authority truthfully or not, for any agent i, reporting its type truthfully is a weakly dominant strategy.

16.2 Relating Mechanism Design to Bayesian Games

Suppose the types are generated from a common prior **P** and type of any player i, θ_i is revealed only to the respective players.
We can put the Mechanism Design scenario in Bayesian Game setting, by considering the following Bayesian Game:
$\langle N, (A_i)_{i \in N}, (\Theta_i)_{i \in N}, P, (\Gamma_\theta)_{\theta \in \Theta} \rangle$ where

- N = Set of Players $\{1, 2, , ..., n\}$
- M_i = Message space of Player i (Corresponding to Action set in Bayesian Games)
- Θ_i = Set of type for player i
- P = Common prior over $\bigtimes_{i \in N} \Theta_i$
- $\Gamma_\theta = \langle N, (A_i)_{i \in N}, (u_i(\theta))_{i \in N} \rangle$

Consider the function $m_i : \Theta_i \to M_i$. We can observe that given any type, the function m_i maps to a member of Message space. It is equivalent to pure strategies of players in a normal Bayesian game. Intuitively every player sends some message to central authority to maximize its utility which are equivalent to strategy a player might have taken in case of bayesian games.

16.3 SCF implememented in Bayesian Equilibrium

Definition 16.2 *A mechanism $\langle M, g\rangle$ implements SCF f in Bayesian Equilibrium if the following two conditions hold:*

1. $\exists(m_1, m_2, ..., m_n)$ *s.t.* $m_i(\theta_i)$ *maximizes the **ex-interim** utility of agent* i, $\forall\theta_i \in \Theta_i, \forall i \in N$, *i.e,*

$$\mathbb{E}_{\theta_{-i}|\theta_i}[u_i(g(m_i(\theta_i), m_{-i}(\theta_{-i})), \theta_i)] \geq \mathbb{E}_{\theta_{-i}|\theta_i}[u_i(g(m_i', m_{-i}(\theta_{-i})), \theta_i)] \qquad \forall\theta_i \in \Theta_i, \forall i \in N \quad (16.2)$$

2. $g(m_i(\theta_i), m_{-i}(\theta_{-i})) = f(\theta_i, \theta_{-i})$

16.3.1 Observation

If SCF f is implementable in Dominant Strategies then f will be implementable in Bayesian Equilibrium.

16.4 Bayesian Incentive Compatibility (BIC)

Definition 16.3 *A direct mechanism $\langle\Theta, f\rangle$ is Bayesian Incentive Compatible if*

$$\mathbb{E}_{\theta_{-i}|\theta_i}[u_i(f(\theta_i, \theta_{-i}), \theta_i)] \geq \mathbb{E}_{\theta_{-i}|\theta_i}[u_i(f(\theta_i', \theta_{-i}), \theta_i)] \qquad \forall\theta_i, \theta_i' \in \Theta_i, \forall i \in N \qquad (16.3)$$

16.4.1 Observation

From equations 16.1 and 16.3, we can observe that the condition for a direct mechanism to be DSIC is required to hold for all $\theta_{-i} \in \Theta_{-i}$ but for it to be BIC this condition is not required because we are doing weighted average over all possible θ_{-i} in some sense.

16.5 Revelation Principle for BI SCFs

Theorem 16.4 *If a Social Choice Function f is implementable in Bayesian Equilibrium then f is Bayesian Incentive Compatible.*

The proof of the above theorem is similar to the proof of Revelation Principle for DSI SCFs and hence left as an exercise.

16.6 Arrovian Social Welfare Function (SWF)

16.6.1 Setup

At this point lets take a step back and ask a question - '*How do you aggregate choices/preferences of different agents?*'
To further understand the implication of this question consider the following scenario -
Suppose you and two of your friends want to watch a movie every day for the next seven days. For this purpose you have chosen seven moves. But each of you want to watch the movie you like the most earlier and each of you have different preference order over the movies. Furthermore the three of you want to watch the movies together. So *How will you aggregate your preferences over all the movies?* so that all three of you can watch the movies together and total satisfaction in watching the movies is maximized.

This question leads us to study of **Arrovian Social Welfare Functions**.

Definition 16.5 *Arrovian Social Welfare Function takes as input individual preferences of different agents outputs a social preference for all the agents.*

16.6.2 Notation

- Set of alternatives $A = \{a_1, ..., a_m\}$.
- Agents $N = \{1, 2, ..., n\}$.
- aR_ib: Alternative a is atleast as good as b for agent i.
- Set of all possible ordering $\mathscr{R}$.

16.6.2.1 Properties of R_i (ordering)

- **Completeness**: For every $a, b \in A$ either aR_ib or bR_ia
- **Reflexivity**: $\forall a \in A$, aR_ia
- **Transitivity**: If aR_ib, bR_ic then aR_ic

Lets divide the relation R_i into two parts P_i(asymmetric part) and I_i(symmetric part) i.e,

- aP_ib: Alternative a is strictly better than b for i.
- aI_ib: Alternative a is indifferent to b for i.

Definition 16.6 *An ordering R_i is* ***linear*** *if aR_ib and bR_ia, then $a = b$ i.e, indifference is not allowed.*

16.6.3 Arrovian Social Welfare Function Representation

Using the notations we have discussed have Arrovian Social Welfare Function F is represented as $F : \mathscr{R}^n \to \mathscr{R}$

Since $F(R)$ is an ordering over the alternatives like R_i we can split it into two parts $\hat{F}(R)$(asymmetric part) and $\overline{F}(R)$(symmetric part) where $R = (R_1, ..., R_n)$.

16.6.4 Desired Properties for Arrovian SWF

Definition 16.7 *A Social Welfare Function F satisfies* ***Weak Pareto*** *if*
$\forall a, b \in A$, $[aP_ib, \forall i \in N] \implies [a\hat{F}(R)b]$

Definition 16.8 *A Social Welfare Function F satisfies* ***Strong Pareto*** *if*
$\forall a, b \in A$, $[aR_ib, \forall i \in N, \exists j, aP_jb] \implies [a\hat{F}(R)b]$

Here we can make an observation that Strong Pareto $\implies$ Weak Pareto.

We want our Arrovian Social Welfare Function to have above two desired Properties.

16.7 Summary

In this lecture we discussed what it means to have SWF f represented in Bayesian Equilibrium, Bayesian Incentive Compatibility. Then we started *Social Choice Theory* by discussing Arrovian Social Welfare Function, its setup, notation and properties. In the next lecture we will continue our discussion on *Arrovian Social Welfare Functions.*