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

# Lecture 20: September 15, 2017

*Lecturer: Swaprava Nath* *Scribe(s): Aayush Ojha*

**Disclaimer**: *These notes aggregate content from several texts and have not been subjected to the usual scrutiny deserved by formal publications. If you find errors, please bring to the notice of the Instructor.*

## 20.1 Recap

In the last lecture we defined Pareto Efficiency, Unanimity and Ontoness for Social choice functions. We also showed that Pareto Efficiency $\implies$ Unanimity $\implies$ Ontoness. Then we defined the Monotone and Strategy-proof Social choice function and showed that they are equivalent. In this lecture we will look at Gibbard-Satterwaithe theorem which states if number of alternatives are more than 2 then any onto and monotone Social choice function will be dictatorial.

## 20.2 Gibbard-Satterthwaite theorem

We define dictatorial Social choice function.

**Definition 20.1** *A Social choice function is* ***dictatorial*** *if it always selects the first preference of a distinguished agent (dictator).*

First we look at a result from last lecture which will be used in showing the Gibbard-Satterthwaite theorem.

**Theorem 20.2** *A SCF, $f$ is Onto and Strategy-Proof $\implies$ $f$ is Unanimous and Strategy-Proof $\implies$ $f$ is Pareto-Efficient and Strategy-Proof.*

We will now formally state the Gibbard-Satterthwaite theorem.

**Theorem 20.3** *Suppose $\mid A \mid \geq 3$. If $f$ is Onto and Strategy-Proof then $f$ is dictatorial.*

Before proving this theorem we will look at some cases where the Gibbard-Satterthwaite theorem may not hold.

### 20.2.1 Cases where Gibbard-Satterthwaite theorem may not hold

#### 20.2.1.1 Restricted Preferences

We assume that all preference profile i.e. all possible order for each agent is possible. In a setting where these preferences are restricted, Gibbard-Satterthwaite theorem may not hold.

#### 20.2.1.2 $\mid \mathcal{A} \mid = 2$

If number of alternatives is two we can construct a Social choice function which is Onto and Strategy-Proof but not dictatorial. Plurality with a fixed tie breaking rule is Strategy-Proof, Onto and non-dictatorial.

#### 20.2.1.3 Indifferences in preferences

If indifferences are allowed among various alternative, then generally Gibbard-Satterthwaite theorem does not hold. But in proof we will use some specific profile constructions. If these profile construction are possible Gibbard-Satterthwaite theorem holds.

Also note that Gibbard-Satterthwaite theorem will hold true even when agents provide real number utilities for each alternative as long as the ordinal order is maintained. Thus Gibbard-Satterthwaite theorem holds on cardinalisation as long as the ordinal order is maintained by utilities.

### 20.2.2 Proof

We will look at proof provided by [Sen01]. For simplicity we will prove the theorem only for the case when number of agents, $n = 2$. Let $\mathcal{N} = \{1, 2\}$.

**Lemma 20.4** *$\mid \mathcal{A} \mid \geq 3$, $\mathcal{N} = \{1, 2\}$, $f$ is Onto and Strategy-Proof then for every preference profile $\mathcal{P}$, $f(\mathcal{P}) \in \{P_1(1), P_2(1)\}$.*

**Proof:** First we look at case where first preference of both agent is same i.e. $P_1(1) = P_2(1)$. As $f$ is unanimous using theorem 20.2, $f(\mathcal{P}) = P_1(1) = P_2(1)$. Let $a = P_1(1) \neq b = P_2(1)$ and $c \in \mathcal{A}$ such that $f(\mathcal{P}) = c \neq a, b$. We create following four preference profiles.

| $P_1$ | $P_2$ | $P_1$ | ${P_2}'$ | ${P_1}'$ | ${P_2}'$ | ${P_1}'$ | $P_2$ |
|---|---|---|---|---|---|---|---|
| $a$ | $b$ | $a$ | $b$ | $a$ | $b$ | $a$ | $b$ |
| - | - | - | $a$ | $b$ | $a$ | $b$ | - |
| - | - | - | - | - | - | - | - |

${P_2}'$ is created by putting $a$ at second preference in $P_2$ and shifting other alternative by 1 place. ${P_1}'$ is created in similar way using $P_1$. $f(P_1, {P_2}') \in \{a, b\}$ because $a$ pareto dominates every other alternative except $b$. If $f(P_1, {P_2}') = b$, agent 2 can manipulate by reporting ${P_2}'$ in place of $P_2$ where preferences for first agent is $P_1$. But $f$ is Strategy-Proof. Thus $f(P_1, {P_2}') = a$. Similarly, we can argue that $f(P_1, {P_2}') = b$.

Now consider transition from $(P_1, {P_2}')$ to $({P_1}', {P_2}')$, position of $a$ improves. Thus using Monotonicity of $f$, $f({P_1}', {P_2}') = a$. Again consider transition from $({P_1}', P_2)$ to $({P_1}', {P_2}')$, position of $b$ improves. Thus using Monotonicity of $f$, $f({P_1}', {P_2}') = b$. Thus we have a contradiction.

This proves that $f(\mathcal{P}) \in \{P_1(1), P_2(1)\}$. ■

Note that the above lemma reduces the values taken any Social choice function $f$ to first preferences of agents. Next lemma will prove the Gibbard-Satterthwaite theorem for 2 agents.

**Lemma 20.5** *$\mid \mathcal{A} \mid \geq 3$, $\mathcal{N} = \{1,2\}$ and $f$ is Onto and Strategy-Proof. Let $\mathcal{P} : P_1(1) = a \neq b = P_2(1)$ and $\mathcal{P}' : {P_1}'(1) = c \neq d = {P_2}'(1)$.*

$$[f(\mathcal{P}) = a] \implies [f(\mathcal{P}') = c]$$

*and*

$$[f(\mathcal{P}) = b] \implies [f(\mathcal{P}') = d]$$

**Proof:** We will first show that $[f(\mathcal{P}) = a] \implies [f(\mathcal{P}') = c]$. For this we will assume that $f(\mathcal{P}) = a$ and show that $f(\mathcal{P}') = c$. To show this we will do a case by case analysis. These cases are:

- $c = a$ and $d = b$
- $c \neq a, b$ and $d = b$
- $c \neq a, b$ and $d \neq b$
- $c = a$ and $d \neq a, b$
- $c = b$ and $d \neq a, b$
- $c = b$ and $d = a$

*Case 1: $c = a$ and $d = b$*
For sake of contradiction we will assume $f(\mathcal{P}') = d = b$. We construct preference profiles as follows.

| $P_1$ | $P_2$ | ${P_1}'$ | ${P_2}'$ | $\widehat{P_1}$ | $\widehat{P_2}$ |
|---|---|---|---|---|---|
| $a$ | $b$ | $a$ | $b$ | $a$ | $b$ |
| - | - | - | - | $b$ | $a$ |
| - | - | - | - | - | - |

Consider transition from $(P_1, P_2)$ to $(\widehat{P_1}, \widehat{P_2})$. Preference for $a$ improves for both agents and $f(P_1, P_2) = a$. Thus by monotonicity, $f(\widehat{P_1}, \widehat{P_2}) = a$. Next consider transition from $({P_1}', {P_2}')$ to $(\widehat{P_1}, \widehat{P_2})$. Preference for $b$ improves for both agents and $f({P_1}', {P_2}') = b$. Thus by monotonicity, $f(\widehat{P_1}, \widehat{P_2}) = b$. But $a \neq b$. This gives us the contradiction. Therefore, $f(\mathcal{P}') = c = a$.

*Case 2: $c \neq a, b$ and $d = b$*
For sake of contradiction we will assume $f(\mathcal{P}') = d = b$. We construct preference profiles as follows.

| $P_1$ | $P_2$ | ${P_1}'$ | ${P_2}'$ | $\widehat{P_1}$ | $P_2$ |
|---|---|---|---|---|---|
| $a$ | $b$ | $c$ | $b$ | $c$ | $b$ |
| - | - | - | - | $a$ | - |
| - | - | - | - | - | - |

First we consider transition from $({P_1}', {P_2}')$ to $(\widehat{P_1}, P_2)$. Notice that this transition is satisfy all constraints of *case 1*. Thus using result from *Case 1*, we get $f(\widehat{P_1}, P_2) = b$.

Consider preference profile $(\widehat{P_1}, P_2)$. $f(\widehat{P_1}, P_2) = b$ and $f(P_1, P_2) = a$. If agent 1 instead of $\widehat{P_1}$ reports $P_1$ he will get a better choice. Thus $f$ is not Strategy-Proof which is a contradiction. Therefore, $f(\mathcal{P}') = c$.

*Case 3: $c \neq a, b$ and $d \neq b$*
For sake of contradiction we will assume $f(\mathcal{P}') = d \neq b$. We construct preference profiles as follows.

| $P_1$ | $P_2$ | ${P_1}'$ | ${P_2}'$ | $\widehat{P_1}$ | $\widehat{P_2}$ |
|---|---|---|---|---|---|
| $a$ | $b$ | $c$ | $d$ | $c$ | $b$ |
| - | - | - | - | $a$ | - |
| - | - | - | - | - | - |

We first consider transition from $({P_1}', {P_2}')$ to $(\widehat{P_1}, \widehat{P_2})$. This transition follows constraints of *case 2*. Thus using result from *case 2* we get $f(\widehat{P_1}, \widehat{P_2}) = b$.

Next, we consider transition from $(P_1, P_2)$ to $(\widehat{P_1}, \widehat{P_2})$. This transition also follows constraints of *case 2*. Thus using result from *case 2* we get $f(\widehat{P_1}, \widehat{P_2}) = c$. But, $b \neq c$. We have a contradiction. Therefore, $f(\mathcal{P}') = c$.

*Case 4: $c = a$ and $d \neq a, b$*
For sake of contradiction we will assume $f(\mathcal{P}') = d \neq a, b$. We construct preference profiles as follows.

| $P_1$ | $P_2$ | ${P_1}'$ | ${P_2}'$ | $\widehat{P_1}$ | $\widehat{P_2}$ |
|---|---|---|---|---|---|
| $a$ | $b$ | $a$ | $d$ | $a$ | $b$ |
| - | - | - | - | - | - |
| - | - | - | - | - | - |

We first consider transition from $({P_1}', {P_2}')$ to $(\widehat{P_1}, \widehat{P_2})$. This transition follows constraints of *case 2* (swap agent 1 and agent 2). Thus using result from *case 2* we get $f(\widehat{P_1}, \widehat{P_2}) = b$.

Next, we consider transition from $(P_1, P_2)$ to $(\widehat{P_1}, \widehat{P_2})$. This transition follows constraints of *case 1*. Thus using result from *case 2* we get $f(\widehat{P_1}, \widehat{P_2}) = a$. But, $b \neq a$. We have a contradiction. Therefore, $f(\mathcal{P}') = a$.

*Case 5: $c = b$ and $d \neq a, b$*
For sake of contradiction we will assume $f(\mathcal{P}') = d \neq a, b$. We construct preference profiles as follows.

| $P_1$ | $P_2$ | ${P_1}'$ | ${P_2}'$ | $\widehat{P_1}$ | $\widehat{P_2}$ |
|---|---|---|---|---|---|
| $a$ | $b$ | $b$ | $d$ | $a$ | $d$ |
| - | - | - | - | - | - |
| - | - | - | - | - | - |

We first consider transition from $({P_1}', {P_2}')$ to $(\widehat{P_1}, \widehat{P_2})$. This transition follows constraints of *case 4* (swap agent 1 and agent 2). Thus using result from *case 4* we get $f(\widehat{P_1}, \widehat{P_2}) = d$.

Next, we consider transition from $(P_1, P_2)$ to $(\widehat{P_1}, \widehat{P_2})$. This transition also follows constraints of *case 4*. Thus using result from *case 4* we get $f(\widehat{P_1}, \widehat{P_2}) = a$. But, $d \neq a$. We have a contradiction. Therefore, $f(\mathcal{P}') = c = b$.

*Case 6: $c = b$ and $d = a$*
For sake of contradiction we will assume $f(\mathcal{P}') = d = a$. We construct preference profiles as follows.

| $P_1$ | $P_2$ | ${P_1}'$ | ${P_2}'$ | $\widehat{P_1}$ | ${P_2}'$ | $\widetilde{P_1}$ | ${P_2}'$ |
|---|---|---|---|---|---|---|---|
| $a$ | $b$ | $b$ | $a$ | $b$ | $a$ | $x$ | $a$ |
| - | - | - | - | $x$ | - | - | - |
| - | - | - | - | - | - | - | - |

We assume $x \neq a, b$. As $\mid \mathcal{A} \mid \geq 3$ such a $x$ will always exists.

We first consider transition from $({P_1}', {P_2}')$ to $(\widehat{P_1}, {P_2}')$. This transition follows constraints of *case 1* (swap agent 1 and agent 2). Thus using result from *case 1* we get $f(\widehat{P_1}, {P_2}') = a$.

Next, we consider transition from $(P_1, P_2)$ to $(\widetilde{P_1}, {P_2}')$. This transition follows constraints of *case 3*. Thus using result from *case 3* we get $f(\widetilde{P_1}, {P_2}') = x$.

Consider preference profile $(\widehat{P_1}, {P_2}')$. $f(\widehat{P_1}, {P_2}') = a$. If instead of $\widehat{P_1}$ agent 1 report its preference as $\widetilde{P_1}$, outcome will be $f(\widetilde{P_1}, {P_2}') = x$. $x$ is a better alternative for agent 1 than $a$. Thus $f$ is not Strategy-Proof which is a contradiction.

Therefore, $f(\mathcal{P}') = c = b$.

This shows that $[f(\mathcal{P}) = a] \implies [f(\mathcal{P}') = c]$. To show that $[f(\mathcal{P}) = b] \implies [f(\mathcal{P}') = d]$ we can use similar arguments.

■

# References

[Sen01] Arunava Sen. Another direct proof of the gibbard-satterthwaite theorem. *Economics Letters*, 70(3):381–385, 2001.