CS698W: 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 at swaprava@cse.iitk.ac.in.*

20.1 Recap

In the last lecture we defined Pareto efficiency (PE), unanimity (UN), and ontoness (ONTO) for social choice functions. We also showed that PE $\implies$ UN $\implies$ ONTO. Then we defined monotonicity and showed that that is equivalent to strategyproofness. In this lecture, we will look at the 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, which is the 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 strategyproof $\Leftrightarrow$ f is Unanimous and strategyproof $\Leftrightarrow$ f is Pareto efficient and strategyproof.*

We will now formally state the theorem.

Theorem 20.3 (Gibbard (1973), Satterthwaite (1975)) *Suppose $|A| \geqslant 3$. If f is onto and strategyproof then f is dictatorial.*

Note that, due to Theorem 20.2, this theorem can equivalently be stated w.r.t. unanimity or Pareto efficiency.

20.2.1 Discussions

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. An example: single-peaked preferences.

20.2.1.2 $|A| = 2$

If number of alternatives is two we can construct a social choice function which is onto and strategyproof but not dictatorial. Plurality with a fixed tie breaking rule is strategyproof, 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. In proof we will use some specific profile constructions. If these profile construction are possible Gibbard-Satterthwaite theorem holds.

20.2.1.4 Cardinalization

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 cardinalization as long as the ordinal order is maintained by utilities.

20.2.2 Proof of Theorem 20.3

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 $N = \{1, 2\}$.

Lemma 20.4 *Let $|A| \geqslant 3$ and $N = \{1, 2\}$. If f is onto and strategyproof then for every preference profile P, $f(P) \in \{P_1(1), P_2(1)\}$.*

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, $f(P) = P_1(1) = P_2(1)$. Let $P_1(1) = a \neq b = P_2(1)$ and $c \in A$ such that $f(P) = c \neq a, b$. 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	-
-	-	-	-	-	-	-	-

The preference $P_2{}'$ is created by putting a at second preference in P_2 and shifting other alternative by at most one place. Similarly, $P_1{}'$ is created in similar way from P_1 by placing b at the second position. Note that, $f(P_1, P_2{}') \in \{a, b\}$ because f is PE (an equivalent condition due to Theorem 20.2) and a Pareto dominates every other alternative except b. But 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. Since f is strategyproof, it implies that $f(P_1, P_2{}') = a$. Similarly, we can argue that $f(P_1, P_2{}') = b$.

Consider the transition from $(P_1, P_2{}')$ to $(P_1{}', P_2{}')$, position of a weakly improves. Thus using monotonicity of f, we conclude $f(P_1{}', P_2{}') = a$. However, if we consider the transition from $(P_1{}', P_2)$ to $(P_1{}', P_2{}')$, position of b weakly improves. Thus by monotonicity of f, we conclude $f(P_1{}', P_2{}') = b$, which is a contradiction. Hence we have the lemma. ■

The above lemma reduces the social choice outcome to the top alternatives of the agents. The next lemma will prove the Gibbard-Satterthwaite theorem for 2 agents.

Lemma 20.5 *Let $|A| \geqslant 3$, $N = \{1, 2\}$, and f is onto and strategyproof. Let $P : P_1(1) = a \neq b = P_2(1)$ and $P' : {P_1}'(1) = c \neq d = {P_2}'(1)$. Then*

$$f(P) = a \implies f(P') = c, \quad \textit{and,}$$
$$f(P) = b \implies f(P') = d.$$

Proof: We will only show that if $f(P) = a$ then $f(P') = c$, since the other case is symmetric and the same proof works.To show this we consider the following exhaustive cases.

1. $c = a$ and $d = b$
2. $c \neq a, b$ and $d = b$
3. $c \neq a, b$ and $d \neq b$
4. $c = a$ and $d \neq a, b$
5. $c = b$ and $d \neq a, b$
6. $c = b$ and $d = a$

Case 1: $c = a$ and $d = b$: Suppose for contradiction $f(P') = d = b$. We construct preference profiles as follows.

P_1	P_2	${P_1}'$	${P_2}'$	$\hat{P}_1$	$\hat{P}_2$
a	b	a	b	a	b
-	-	-	-	b	a
-	-	-	-	-	-

Consider transition from (P_1, P_2) to $(\hat{P}_1, \hat{P}_2)$. Preference for a improves for both agents and $f(P_1, P_2) = a$. Thus by monotonicity, $f(\hat{P}_1, \hat{P}_2) = a$. Next consider transition from $({P_1}', {P_2}')$ to $(\hat{P}_1, \hat{P}_2)$. Preference for b improves for both agents and $f({P_1}', {P_2}') = b$. Thus by monotonicity, $f(\hat{P}_1, \hat{P}_2) = b$. But $a \neq b$. This gives us a contradiction. Therefore, $f(P') = c = a$.

Case 2: $c \neq a, b$ and $d = b$: Suppose for contradiction $f(P') = d = b$. We construct preference profiles as follows.

P_1	P_2	${P_1}'$	${P_2}'$	$\hat{P}_1$	P_2
a	b	c	b	c	b
-	-	-	-	a	-
-	-	-	-	-	-

First we consider transition from $({P_1}', {P_2}')$ to $(\hat{P}_1, P_2)$. Notice that this transition satisfies all constraints of *Case 1*. Hence, $f(\hat{P}_1, P_2) = b$.

Consider preference profile $(\hat{P}_1, P_2)$. At this profile if agent 1 reports P_1 instead of $\hat{P}_1$, the outcome is a which she prefers more than the current outcome b, as $f(\hat{P}_1, P_2) = b$ and $f(P_1, P_2) = a$. This is a contradiction to f being strategyproof. Therefore, $f(P') = c$.

Case 3: $c \neq a, b$ and $d \neq b$: Suppose for contradiction $f(P') = d \neq b$. We construct preference profiles as follows.

P_1	P_2	${P_1}'$	${P_2}'$	$\hat{P}_1$	$\hat{P}_2$
a	b	c	d	c	b
-	-	-	-	a	-
-	-	-	-	-	-

We first consider transition from $({P_1}', {P_2}')$ to $(\hat{P}_1, \hat{P}_2)$. This transition follows the constraints of *Case 2*. Hence, $f(\hat{P}_1, \hat{P}_2) = b$.

Next, we consider transition from (P_1, P_2) to $(\hat{P}_1, \hat{P}_2)$. This transition also follows the constraints of *Case 2*. Hence $f(\hat{P}_1, \hat{P}_2) = c$. But, $b \neq c$. We have a contradiction. Therefore, $f(P') = c$.

Case 4: $c = a$ and $d \neq a, b$: Suppose for contradiction $f(P') = d \neq a, b$. We construct preference profiles as follows.

P_1	P_2	${P_1}'$	${P_2}'$	$\hat{P}_1$	$\hat{P}_2$
a	b	a	d	a	b
-	-	-	-	-	-
-	-	-	-	-	-

We first consider transition from $({P_1}', {P_2}')$ to $(\hat{P}_1, \hat{P}_2)$. This transition follows constraints of *Case 2* (swap agent 1 and agent 2). Hence $f(\hat{P}_1, \hat{P}_2) = b$.

Next, we consider transition from (P_1, P_2) to $(\hat{P}_1, \hat{P}_2)$. This transition follows constraints of *Case 1*. Hence $f(\hat{P}_1, \hat{P}_2) = a$. But, $b \neq a$. We have a contradiction. Therefore, $f(P') = a$.

Case 5: $c = b$ and $d \neq a, b$: Suppose for contradiction $f(P') = d \neq a, b$. We construct preference profiles as follows.

P_1	P_2	${P_1}'$	${P_2}'$	$\hat{P}_1$	$\hat{P}_2$
a	b	b	d	a	d
-	-	-	-	-	-
-	-	-	-	-	-

We first consider transition from $({P_1}', {P_2}')$ to $(\hat{P}_1, \hat{P}_2)$. This transition follows constraints of *Case 4* (swap agent 1 and agent 2). Hence $f(\hat{P}_1, \hat{P}_2) = d$.

Next, we consider transition from (P_1, P_2) to $(\hat{P}_1, \hat{P}_2)$. This transition also follows constraints of *Case 4*. Hence $f(\hat{P}_1, \hat{P}_2) = a$. But, $d \neq a$. We have a contradiction. Therefore, $f(P') = c = b$.

Case 6: $c = b$ and $d = a$: Suppose for contradiction $f(P') = d = a$. We construct preference profiles as follows.

P_1	P_2	${P_1}'$	${P_2}'$	$\hat{P}_1$	${P_2}'$	$\tilde{P}_1$	${P_2}'$
a	b	b	a	b	a	x	a
-	-	-	-	x	-	-	-
-	-	-	-	-	-	-	-

We assume $x \neq a, b$. As $|A| \geqslant 3$ such a x will always exist.

We first consider transition from $({P_1}', {P_2}')$ to $(\hat{P}_1, {P_2}')$. This transition follows constraints of *Case 1* (swap agent 1 and agent 2). Hence $f(\hat{P}_1, {P_2}') = a$.

Next, we consider transition from (P_1, P_2) to $(\tilde{P}_1, {P_2}')$. This transition follows constraints of *Case 3*. Hence $f(\tilde{P}_1, {P_2}') = x$.

Consider preference profile $(\hat{P}_1, {P_2}')$. At this profile $f(\hat{P}_1, {P_2}') = a$. If instead of $\hat{P}_1$ agent 1 report its preference to be $\tilde{P}_1$, outcome will be $f(\tilde{P}_1, {P_2}') = x$, which is more preferred by agent 1 than a at $\hat{P}_1$. This is a contradiction to f being strategyproof. Therefore, $f(P') = c = b$.

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

The remaining part of the proof (for more than two agents) uses induction on the number of agents and is skipped here. An interested reader may look up [Sen01]. ■

References

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