Lecture 10

Lecturer: Ariel Procaccia *Scribe: Stylianos Despotakis*

1 Overview

In the previous lecture, we discussed how we can use complexity as a barrier to avoid the manipulation of a voting rule. In this lecture, we continue by proving the theorem about the complexity of "score-based" rules. Also, we discuss why the NP-hardness of manipulation problem is not enough to say that a voting rule is "nonmanipulable", and we see some different approaches to studying manipulation.

2 Complexity of Manipulation

2.1 The R-Manipulation Problem

Let R be a voting rule and suppose that a manipulator has a preferred candidate p. Assume also that the manipulator knows the votes of every other voter. Is it possible for the manipulator to vote in a way that makes p the unique winner under R? A natural algorithm for the manipulator to achieve this, is the following greedy algorithm:

1. Rank p in the first place.
2. While there are unranked alternatives:
   - If there is an alternative that can be placed in the next spot without preventing p from winning, place this alternative in the next spot and continue.
   - Otherwise, return "false".

If the above algorithm returns a ranking, then the manipulator can vote according to this ranking and make p the winner. If, on the other hand, the algorithm returns "false", this means that it couldn't find such a ranking. However, that doesn't necessarily means that such a ranking doesn't exist, only that this algorithm couldn't find one.

The following theorem defines a class of rules for which the above algorithm always decides R-Manipulation correctly, i.e. if the algorithm returns false, there is no ranking which makes p the winner.

Theorem 1 (Bartholdi et al., SCW 89) *Fix $i \in N$ and the votes of other voters. Let R be a rule such that there exists a function $s(\prec_i, x)$ such that:*

- *For every $\prec_i$ chooses a candidate that uniquely maximizes $s(\prec_i, x)$.*
- *For every $\prec_i, \prec_i'$ such that $\{y : y \prec_i x\} \subseteq \{y : y \prec_i' x\}$, it holds that $s(\prec_i, x) \leq s(\prec_i', x)$.*

Then the greedy algorithm always decides R-Manipulation correctly.

Proof: Suppose that the algorithm failed to return a ranking and assume, for contradiction, that there exists a $\prec'$ that makes p the winner. Let $\prec$ be the unfinished ranking (output) of the algorithm. Let U be the set of alternatives not ranked by $\prec$, and u be the highest ranked alternative in U according to $\prec'$. We complete $\prec$ by adding u in the first available spot and the other candidates in $U \setminus \{u\}$ arbitrarily.

Example: $A = \{p, a, b, c, d\}$

	$\prec$	$\prec'$	
output of alg.	p	b	
output of alg.	b	p	
output of alg.	d	ⓐ	
	$a = u$	d	$U = \{a, c\}$
	c	ⓒ	

Since p is above more candidates in $\prec$ than in $\prec'$, from the second property of s we derive that $s(\prec, p) \geq s(\prec', p)$. Moreover, from the first property we know that p uniquely maximizes $s(\prec', x)$, since $\prec'$ makes p the winner under R. Therefore, $s(\prec', p) > s(\prec', u)$. Finally, from the second property again, we have that $s(\prec', u) \geq s(\prec, u)$.

Thus, $s(\prec, p) > s(\prec, u)$. In other words the algorithm could put u in the next position without preventing p from winning, which is a contradiction. ■

One example of a scoring rule that satisfies the properties in the theorem is Plurality. Let $a_1, a_2, \ldots, a_m$ be a fixed order of the candidates, and assume that in case of a tie among some candidates, the first of them according to this ordering is the winner. We can define the function s as

$$s(\prec_i, a_k) = (\#\text{ of times that } a_k \text{ is first}) + \frac{m-k}{m}.$$

We can easily check that this function satisfies the two properties.

2.2 Criticisms

There are some criticisms for the approach of manipulation under the complexity perspective. For example, the Dictatorship-Manipulation problem is easy: if you are the dictator you put the preferred candidate first, otherwise you can't manipulate. However, Dictatorship is not considered a "manipulable" system.

Moreover, there are some rules for which the corresponding problem is NP-hard in the worst case, but a manipulator can usually succeed. In fact, most reasonable voting rules are usually manipulable.

There are two alternative approaches:

- Algorithmic: this approach is for specific voting rules but works for every reasonable distribution.
- Quantitative G-S: this approach is for every reasonable voting rule but works for a specific distribution (it is difficult to extend the results for every distribution).

Below, we will focus on Quantitative G-S approach.

3 Quantitative G-S

Definition 2 *We define the distance between two voting rules f, g to be the fraction of inputs on which they differ, i.e.*

$$D(f, g) = \Pr[\mathrm{f}(\prec) \neq \mathrm{g}(\prec)],$$

where the probability is over uniformly random preference profiles $\prec$.

Also, we define the distance of a voting rule f to a set of voting rules F as usual, i.e. the minimum distance of f to a rule in F.

Let F_{dic} be the set of dictatorships. Since there are n voters, we have that $|F_{dic}| = n$.

Definition 3 *We say that $(\succ, \succ_i')$ is a manipulation pair for f if $f(\succ_i', \succ_{-i}) \succ_i f(\succ)$, i.e. if voter i change from $\succ_i$ to $\succ_i'$, she can achieve better outcome.*

Theorem 4 (Mossel and Racz, 2012) *Let $m \geq 3$, f be an onto rule, and $D(f, F_{dic}) \geq \epsilon$. Then*

$$\Pr[(\succ, \succ_{\mathrm{i}}') \textit{ is a manipulation pair for } \mathrm{f}] \geq \mathrm{p}\left(\epsilon, \frac{1}{\mathrm{n}}, \frac{1}{\mathrm{m}}\right),$$

for a polynomial p, where $\succ$ and $\succ_i'$ are chosen uniformly at random.

Note that Gibbard-Satterthwaite theorem is a special case of the above theorem: if $m \geq 3$,f is onto, and f is not a dictatorship ($\epsilon > 0$), then f is not strategy-proof ($\Pr[\ldots] > 0$).

Every reasonable voting rule has $D(f, F_{dic}) \geq \epsilon$ for a constant ϵ. Therefore, the above theorem says that for every reasonable voting rule and for a non-negligible fraction of preference profiles we can find a manipulation in polynomial time simply by drawing random votes.

4 Randomized Voting Rules

4.1 Strategy-proof randomized rules

A different approach to manipulation is to consider randomized rules instead of deterministic. A randomized voting rule is a rule which outputs a distribution over the alternatives, instead of a single alternative.

Several questions arise from this definition. For example, what does manipulation mean for a randomized rule? How can we compare two distributions? An elegant approach is to consider the existence of utility functions u_i for each voter i. We assume also that each voter has strict preferences, i.e. there are no two candidates whom she prefers the same.

Definition 5 *We say that $\succ_i$ is consistent with the utility function u_i, if*

$$x \succ_i y \iff u_i(x) > u_i(y).$$

Definition 6 *We say that a randomized voting rule f is strategy-proof, if $\forall i \in N$, $\forall u_i$, $\forall \succ_{-i}$, and $\forall \succ'_i$, it holds that*

$$\mathbb{E}[u_i(f(\prec))] \geq \mathbb{E}[u_i(f(\prec'_i, \prec_{-i}))]$$

where $\succ_i$ is consistent with u_i.

In other words, in the definition above we consider that the expectation reflects the real preference.

Now consider the following two simple types of (deterministic) rules:

- Unilateral rules: these are the rules for which the result depends only on one voter (e.g. a dictatorship, or a rule for which we restricts the range to $\{a, b, c\}$ and choose what a fixed i prefers from these three alternatives, and so on).
- Duple rules: these are the rules with range of size at most 2 (i.e. we fix two alternatives $\{a, b\}$ and we choose one of this according to some rule).

Definition 7 *A randomized rule is called a probability mixture over (deterministic) rules* $f_1, f_2, \ldots, f_k$*, if there exist* $a_1, a_2 \ldots, a_k$ *such that for all* $\succ$, $\Pr[f(\succ) = f_j(\succ)] = a_j$.

The following theorem is a generalization of Gibbard-Satterthwaite theorem, for randomized rules.

Theorem 8 (Gibbard, 1977) *A randomized rule is strategy-proof only if it is a probability mixture over unilaterals and duples.*

The reverse of the theorem is not always true. For example consider the (unilateral) rule that for a fixed voter chooses her least preferred candidate. This is not strategy-proof. However, there is an extension of the theorem in which under some specific properties of unilaterals and duples, we can replace the "only if" part with "iff".

Some examples of strategy-proof randomized rules are the following:

- Randomized dictatorship: choose a voter uniformly at random and choose her first candidate.
- Choose a pair of candidates at random and choose the most preferred among these two.

4.2 Randomization and Approximation

Due to Gibbard-Satterthwaite theorem, we know that most popular deterministic rules are not strategy-proof. Therefore a natural question is if we can find strategy-proof randomized rules which "approximate" popular rules. In other words, can we create a strategy-proof randomized rule with similar behavior with some of the popular (deterministic) rules?

Definition 9 *Fix a rule which has a clear notion of score (e.g. Borda) denoted by* $sc(\succ, x)$*. We say that the randomized rule* f *is a* c*-approximation if for every preference profile* $\succ$,

$$\frac{\mathbb{E}[sc(\succ, f(\succ))]}{\max_{x \in A} sc(\succ, x)} \geq c.$$

Consider the simple randomized rule f which chooses an alternative uniformly at random. This rule gives a $\frac{1}{2}$-approximation for Borda, because

$$\mathbb{E}[sc(\succ, f(\succ))] = \sum_{x \in A} \frac{1}{m} sc(\succ, x) = \frac{1}{m} \cdot \frac{nm(m-1)}{2} = \frac{n(m-1)}{2} \geq \frac{\max_{x \in A} sc(\succ, x)}{2}.$$

The following theorem says that we cannot do something better than that.

Theorem 10 (Procaccia, 2010) *No strategy-proof randomized voting rule can approximate Borda to a factor of* $\frac{1}{2} + \omega\left(\frac{1}{\sqrt{m}}\right)$.

Sketch of Proof: The main idea for the proof of the theorem is to apply Yao's minimax principle. Consider the two-player zero-sum game determined by a matrix like the following

	$\succ^1$	...	...	...	...	$\succ^t$
U_1	$\frac{1}{15}$	...	...	...	...	$\frac{2}{21}$
...	...	...	...	...	...	...
U_k	$\frac{7}{15}$	...	...	...	...	$\frac{5}{21}$
D_1	$\frac{4}{15}$	...	...	...	...	$\frac{8}{21}$
...	...	...	...	...	...	...
D_s	$\frac{13}{15}$	...	...	...	...	$\frac{17}{21}$

The set of strategies for column player are all possible preference profiles $\succ^1, \ldots, \succ^t$. The set of strategies for row player are all possible unilateral rules $U_1, \ldots U_k$ and duple rules $D_1, \ldots, D_s$. The number in an entry $(F, \succ^j)$ of the matrix is the ratio of the Borda score of the winner under rule F and preference profile $\succ^j$, over the maximum Borda score under the preference profile $\succ^j$, i.e. $\frac{sc(\succ^j, F(\succ^j))}{\max_{x \in A} sc(\succ^j, x)}$.

A mixed strategy for row player is a distribution over all unilateral and duple rules. Therefore, for a fixed mixed strategy f of the row player and a fixed column $\succ^j$, the expected utility for row player is $\frac{\mathbb{E}[sc(\succ^j, f(\succ^j))]}{\max_{x \in A} sc(\succ^j, x)}$.

On the other hand, a mixed strategy for column player is a distribution over all possible preference profiles. Thus, for a fixed mixed strategy $\succ$ of the column player and a fixed unilateral or duple rule F, the expected utility for column player is $\mathbb{E}\left[\frac{sc(\succ, F(\succ))}{\max_{x \in A} sc(\succ, x)}\right]$.

By minimax theorem, we derive that

$$\max_{f} \min_{\succ^j} \frac{\mathbb{E}[sc(\succ^j, f(\succ^j))]}{\max_{x \in A} sc(\succ^j, x)} = \min_{\succ} \max_{F} \mathbb{E}\left[\frac{sc(\succ, F(\succ))}{\max_{x \in A} sc(\succ, x)}\right],$$

i.e. the expected ratio of the best distribution over unilateral rules and duples against the worst preference profile is equal to the expected ratio of the worst distribution over profiles against the best unilateral rule or duple.

Therefore, any distribution over profiles against the best unilateral or duple will give an upper bound on the approximation ratio of the best distribution over unilateral rules and duples.

Since any strategy-proof randomized rule is a probability mixture over unilaterals and duples, the above upper bound will give also an upper bound for the approximation ratio of the best strategy-proof randomized rule.

Thus, it remains to find an example of a bad distribution which gives the desired upper bound. Consider the distribution corresponding to the following process:

- Choose an $x^* \in A$ uniformly at random.
- Each voter i chooses a random number $k_i \in \{1, \ldots, \sqrt{m}\}$ and puts x^* in position k_i in her vote.
- The other alternatives are ranked cyclically.

Example: $A = \{a, b, c, d\}, x^* = b, k_1 = 2, k_2 = 1, k_3 = 2$

1	2	3
c	**b**	d
b	a	**b**
a	d	c
d	c	a

Since x^* is among the first $\sqrt{m}$ candidates for every voter, the score of x^* is always high. More formally, $sc(\succ, x^*) \geq n(m - \sqrt{m})$. Moreover, every other alternative $x \in A \setminus \{x^*\}$ has score around the average, i.e. $sc(\succ, x) \sim \frac{n(m-1)}{2}$.

Therefore, in order for a rule to get a good approximation, it has to "find" x^*. However, a unilateral rule cannot do that, since by looking at only one vote, there is no way to tell who x^* is among the first $\sqrt{m}$ candidates. Similarly, it is impossible for a duple rule to "find" x^*, since by fixing only two alternatives, the probability that x^* is among them is $\frac{2}{m}$.

Thus, the above distribution is a bad distribution, and it turns out that it gives the desired upper bound. ■