Computational Social Choice: Autumn 2010

Ulle Endriss
Institute for Logic, Language and Computation
University of Amsterdam

Plan for Today

Today we will briefly touch on a number of additional topics in voting theory we did not have time to cover in depth:

- Weighted Voting Games
- Proportional Representation
- Electronic Voting
- various other topics (one slide each)

Example

Consider a parliament in which

- 45% of the MPs belong to Party A,
- 40% of the MPs belong to Party B, and
- 15% of the MPs belong to Party C.

To pass a law, you need the support of $> 50\%$ of the MPs.

Does Party A command more *power* than Party C?

Weighted Voting Games

A *weighted voting game* for a set of voters $\mathcal{N} = \{1, \ldots, n\}$ consists of

- a vector of (nonnegative) *weights* $(w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$ and
- a (positive) *quota* $q \in \mathbb{R}$ (usually at most $w_1 + \cdots + w_n$).

A *coalition* $S \subseteq \mathcal{N}$ is called a *winning coalition* *iff*

$$\sum_{i \in S} w_i \quad \geqslant \quad q$$

Example: Our example on the previous slide corresponds to the weighted voting game with weights $(45, 40, 15)$ and quota 51.

Remark: A weighted voting game is a special case of a (simple) coalitional game. [$v(S) = 1$ if $\sum_{i \in S} w_i \geqslant q$ and $v(S) = 0$ otherwise]

The Banzhaf Power Index

How can we measure the power of voter i in game $G = [q; w_1, \ldots, w_n]$?

Define the set of coalitions for which voter i is *critical*, i.e., S alone is losing but $S \cup \{i\}$ is winning:

$$\mathrm{Crit}_i(G) \quad := \quad \{S \subseteq \mathcal{N} \setminus \{i\} \mid \sum_{j \in S \cup \{i\}} w_j \geqslant q \text{ and } \sum_{j \in S} w_j < q\}$$

The proportion of coalitions (excluding i) for which i is critical is called the (raw) *Banzhaf index* of i:

$$\beta_i(G) \quad := \quad \frac{\#\mathrm{Crit}_i(G)}{2^{n-1}}$$

Note: As the sum of all β_i need not be 1, some authors prefer the term "Banzhaf measure" and reserve "Banzhaf index" for $\beta_i / \sum_{j \in \mathcal{N}} \beta_j$.

J.F. Banzhaf. Weighted Voting Doesn't Work: A Mathematical Analysis. *Rutgers Law Review*, 19:317–343, 1965.

D.S. Felsenthal and M. Machover. *The Measurement of Voting Power*. Edward Elgar: Cheltenham, UK, 1998.

The Shapley-Shubik Power Index

Now suppose our voters enter the room in some order. How likely is it that a coalition makes the quota just as i enters?

The *Shapley-Shubik index* of voter i in game G if defined as:

$$\varphi_i(G) \quad := \quad \frac{1}{n!} \sum_{S \in \mathrm{Crit}_i(G)} |S|! \cdot (n-1-|S|)!$$

(For any group S, there are $|S|! \cdot (n-1-|S|)!$ ways in which first the members of S enter the room, then i, and then everyone else.)

L.S. Shapley and M. Shubik. A Method for Evaluating the Distribution of Power in a Committee System. *American Political Science Review*, 48(3):787–792, 1954.

D.S. Felsenthal and M. Machover. *The Measurement of Voting Power*. Edward Elgar: Cheltenham, UK, 1998.

Complexity

Computing either one of the two power indices is intractable:

Theorem 1 (Matsui and Matsui, 2001) *For both the Banzhaf and the Shapley-Shubik index, deciding whether the index of a given voter exceeds 0 is NP-complete*

Proof: First, observe that $\beta_i(G) > 0$ iff $\varphi_i(G) > 0$ iff $\mathrm{Crit}_i(G) \neq \emptyset$. So we need to prove NP-completeness of deciding, given i, whether there exists a coalition S such that i is critical for S:

- NP-membership: clear (S is the certificate) ✓
- NP-hardness: by reduction from PARTITION (next slide)

Y. Matsui and T. Matsui. NP-Completeness for Calculating Power Indices of Weighted Majority Games. *Theoret. Comp. Sci.*, 263(1–2):305–310, 2001.

Proof

Recall the NP-complete Partition problem:

Partition
Instance: $(w_1, \ldots, w_n) \in \mathbb{N}^n$
Question: Is there a set $I \subseteq \{1, \ldots, n\}$ s.t. $\sum_{i \in I} w_i = \frac{1}{2} \sum_{i=1}^{n} w_i$?

Given an instance of Partition, we build a weighted voting game:

- There are $n + 1$ voters. The first n voters have weights $(w_1, \ldots, w_n)$ and the last voter has weight 1.
- The quota is $q := 1 + \frac{1}{2} \sum_{i=1}^{n} w_i$.

Then the last voter is critical only for a coalition corresponding to weights adding up to exactly $q - 1$, i.e., she can only be critical if the answer to the original Partition problem is YES. ✓

The Core

Power indices are about *fairness*. The next concept is about *stability*.

Suppose a winning coalition generates *value* 1; a losing coalition generates *value* 0. The *grand coalition* $\mathcal{N}$ is winning (by definition).

An *imputation* is specification of the division of the value of the grand coalition: a vector $(p_1, \ldots, p_n) \in \mathbb{R}^n$ with $p_i \geqslant 0$ and $\sum_{i \in \mathcal{N}} p_i = 1$.

An imputation $(p_1, \ldots, p_n)$ is said to be *in the core* if

$$\sum_{i \in S} p_i = 1 \text{ for every winning coalition } S \subseteq \mathcal{N}$$

<u>Discussion:</u> Having a nonempty core is desirable; it means that we can arrange payments in such a manner that no coalition has an incentive to break away from the grand coalition.

Complexity

Theorem 2 (Elkind et al, 2009) *For weighted voting games, nonemptiness of the core can be decided in polynomial time.*

Proof: Call $i \in \mathcal{N}$ a *veto voter* if $i \in S$ whenever S is winning.
Observe that the core is nonempty *iff* there exists a veto voter:

- Suppose i is a veto voter. Then $(p_1, \ldots, p_n)$ with $p_i = 1$ and $p_j = 0$ for $j \neq i$ is in the core. ✓
- Take any $(p_1, \ldots, p_n)$. W.l.o.g., let $p_n > 0$. Then $\sum_{\mathcal{N} \setminus \{n\}} p_i < 1$.
  Suppose there is no veto voter. In particular, n is not a veto voter.
  Thus, coalition $\mathcal{N} \setminus \{n\}$ *is winning*. So $(p_1, \ldots, p_n)$ is *not* in the core. ✓

Hence, to check nonemptiness of the core, we only have to check for each voter whether it is a veto voter. But (by monotonicity) i is a veto voter *iff* $\mathcal{N} \setminus \{i\}$ is not winning. This can be checked in polynomial time. ✓

E. Elkind, L.A. Goldberg, P.W. Goldberg, and M. Wooldridge. On the Computational Complexity of Weighted Voting Games. *Annals of Mathematics and Artificial Intelligence*, 56(2):109–131, 2009.

Proportional Representation

Suppose we are voting for political parties (using the plurality rule). Suppose there are 100 voters and 10 seats.

Party A:	47 votes
Party B:	29 votes
Party C:	24 votes

How many seats should each party get?

This is the problem of *proportional representation*. It is (roughly) equivalent to the problem of *apportionment:* in a federal system, how many seats in the house of representatives should go to each state, given its population?

This topic has not yet received much (any?) attention in the COMSOC community. On the following slides, we review some of the classical procedures and discuss a few of their properties.

M.L. Balinski and H.P. Young. *Fair Representation: Meeting the Ideal of One Man, One Vote*. 2nd edition, Bookings Institution Press, 2001.

Hamilton's Method

In the context of assigning seats in the US Congress to states, Alexander Hamilton proposed the following method in 1792:

- Compute the *quota* for each party i:

$$q_i \ := \ \frac{\#\text{votes for } i}{\#\text{votes in total}} \times \#\text{seats}$$

- To each party i, award (for now) $\lfloor q_i \rfloor$ seats.
- Award remaining seats to those parties with the largest fractions.

<u>Remark:</u> The last step may require tie-breaking.

The Alabama Paradox

Suppose there are 250 voters. Consider the outcome under Hamilton's Method when there are 25 seats *vs.* when there are 26 seats:

	votes	votes/25	seats/25	votes/26	seats/26
Party A:	24	2.400	3	2.496	2
Party B:	113	11.300	11	11.752	12
Party C:	113	11.300	11	11.752	12

That is, even though the total number of seats increases, the number of seats for Party A decreases.

This paradox was observed in 1880 in the US when Congress had to fix the number of representatives based on the latest census data: Alabama would get 8 representatives out of 299 but only 7 out of 300.

Jefferson's Method

Let s be the number of seats to be allocated. Let p be the number of parties and let n_i be the number of votes received by party $i \leqslant p$.

Also in 1792, Thomas Jefferson proposed this method:

- Fix a divisor d such that

$$\lfloor n_1/d \rfloor + \lfloor n_2/d \rfloor + \cdots + \lfloor n_p/d \rfloor \quad = \quad s$$

- Award $\lfloor n_i/d \rfloor$ seats to party i.

Observation 1: The number of seats assigned to each party increases monotonically with the number of total seats, so Jefferson's Method does not suffer from the Alabama Paradox.

Observation 2: Jefferson's Method tends to favour larger parties.

Webster's Method

Let $\mathrm{round}(x) := \lfloor x + 0.5 \rfloor$.

In 1832, Daniel Webster proposed this variant of Jefferson's Method:

- Fix a divisor d such that

$$\mathrm{round}(n_1/d) + \mathrm{round}(n_2/d) + \cdots + \mathrm{round}(n_p/d) \quad = \quad s$$

- Award $\mathrm{round}(n_i/d)$ seats to party i.

Electronic Voting

Maybe the most obvious application of techniques from computer science to voting is *electronic voting*.

- Narrowly interpreted, this is about the design of suitable electronic voting machines (i.e., computers) to record and aggregate ballots.
- More generally, research on electronic voting encompasses all aspects of applying concepts from information security research (in particular, cryptography) to voting.

Remark: In the Netherlands, voting by means of electronic voting machines has been abolished in 2008, after a major scandal.

B. Jacobs and W. Pieters. Electronic Voting in the Netherlands: From Early Adoption to Early Abolishment. In *Foundations of Security Analysis and Design V*, Springer-Verlag, 2009.

Verifiability and Privacy

Research into electronic voting has highlighted the following conflicting demands (which are also problematic for traditional elections);

- On the one hand, we want the election result to be *verifiable:*
  - Anybody should be able to do a recount of the ballots. [possible in principle in traditional elections]
  - Each voter should be able to check that her ballot got counted. [impossible in traditional elections]
- The *privacy* of the voter should be guaranteed:
  - Each voter should be able to keep her vote secret. [possible in traditional elections, unless many officials conspire]
  - Even if she wants to, a voter should be unable to prove to others how she voted (to protect against bribery etc.). [possible in traditional elections, modulo camera phones etc.]

Most work aimed at guaranteeing both uses cryptographic methods.

ThreeBallot Voting

ThreeBallot Voting is an interesting proposal by Rivest (2006) that does not rely on cryptography (though it has some known weaknesses).

Suppose we want to elect a single alternative using plurality.

At the election:

- Each voter gets *three ballot sheets*, with different *serial numbers*. The serial numbers are assumed to be hard to remember.
- To *vote*, first mark each alternative on exactly one of the three sheets. Then mark the one you want to vote for on a second sheet.
- Use a (trusted) machine to *check* that you have filled in your triplet in a valid manner (no over-voting, etc.).
- Ask the (trusted) machine to *copy* one of your sheets for you as a take-home receipt. Put all three originals in the ballot box.

R. Rivest. The ThreeBallot Voting System. Technical Report, MIT-CSAIL, 2006.

ThreeBallot Voting (cont.)

After the election:

- All ballots get published on a website (with their serial numbers).
- The alternative with the most votes wins (note that this is like plurality, except that each alternative gets an extra n points).

As a voter, you can

- verify that the ballot you copied has been counted correctly. (A possible attacker does not know which of your three ballot sheets you copied.)
- verify the ballots have been tallied correctly.

As a voter, you cannot

- prove to anybody how you voted (so you cannot be bribed).

Attacks

Unfortunately, there are some problems with ThreeBallot Voting:

- You could bribe a voter to vote using a fixed pattern across all three ballots and only pay if all three types of ballots show up on the website. If the probability of the pattern showing up by chance is small, then this will work with high probability.
- You could bribe voters to give you a receipt with a certain pattern, and then commit fraud on other ballots (which are then less likely to those for which voters have receipts).
- Other weak points are the serial numbers (if you can remember them, you can sell your vote) and the checking/copying machine.

Even More Topics in Voting

On the following slides we will briefly list a few additional topics in voting theory that are relevant to computational social choice and mention a couple of typical references for each of them.

Tournament Solutions

Given a set of voters with linear ballots over a set of alternatives $\mathcal{X}$, we can generate the *majority graph* on $\mathcal{X}$: there is an edge from node $x \in \mathcal{X}$ to node $y \in \mathcal{X}$ iff a majority of voters rank x above y.

If the number of voters is odd, then the majority graph is complete.

A complete and asymmetric relation on $\mathcal{X}$ is also called a *tournament*.

Many voting procedures (e.g., Banks and Copeland), and more generally "solution concepts", can be defined on tournaments.

J.-F. Laslier. *Tournament Solutions and Majority Voting*. Springer-Verlag, 1997.

F. Brandt, F. Fischer, P. Harrenstein, and M. Mair. A Computational Analysis of the Tournament Equilibrium Set. *Soc. Choice Welf.*, 34(4):597–609, 2010.

Algorithm Design for Intractable Voting Procedures

We have seen a number of voting procedures for which the winner determination problem is NP-hard (e.g., Dodgson or Kemeny).

An important line of work is aimed at developing algorithms for these intractable procedures that will perform well in practice.

Of course, algorithm design is also relevant for other hard problem arising in the context of voting, such as the possible winner problem.

A. Davenport and J. Kalagnanam. A Computational Study of the Kemeny Rule for Preference Aggregation. Proc. AAAI-2004.

V. Conitzer. Computing Slater Rankings Using Similarities among Candidates. Proc. AAAI-2006.

Approximation

If winner determination is intractable, it may be more feasible to develop algorithms to compute approximate winners.

Example: We may succeed in developing an algorithm that efficiently computes, say, the Dodgson score of an alternative with a certain (guaranteed) level of precision.

These approximation algorithms themselves may again be viewed as voting procedures, and can be analysed using the tools of social choice theory (e.g., an approximate version of a voting procedure may satisfy an attractive axiom violated by the original procedure).

J.C. McCabe-Dansted, G. Pritchard, and A. Slinko. Approximability of Dodgson's Rule. *Social Choice and Welfare*, 31(2):311–330, 2008.

I. Caragiannis, C. Kaklamanis, N. Karanikolas, and A.D. Procaccia. Socially Desirable Approximations for Dodgson's Voting Rule. Proc. EC-2010.

Parametrised Complexity

Standard complexity analysis (which we have applied to various problems in voting) can be somewhat crude as it does not allow us to pinpoint which parameter of the problem is chiefly responsible for an explosion in complexity.

The theory of *parametrised complexity* allows for a more detailed analysis. It has been applied to a range of problems in social choice theory, particularly in voting.

R.G. Downey and M.R. Fellows. *Parameterized Complexity*. Springer-Verlag, 1999.

N. Betzler, J. Guo, and R. Niedermeier. Parameterized Computational Complexity of Dodgson and Young Elections. *Inform. Comput.*, 208(2):165–177, 2010.

Dynamics of Repeated Voting

Consider a situation where we hold a sequence of elections on the same issue and voters can change their ballot in response to what they have learned during the previous election (e.g., about other voters' choices or about the outcome).

A small number of papers has been concerned with the analysis of the dynamics of such repeated voting games.

Remark: Note that this is different from *sequential voting* as discussed in the lecture on voting in combinatorial domains.

S. Airiau and U. Endriss. Iterated Majority Voting. Proc. 1st International Conference on Algorithmic Decision Theory (ADT-2009).

R. Meir, M. Polukarov, J.S. Rosenschein and N.R. Jennings. Convergence to Equilibria in Plurality Voting. Proc. AAAI-2010.

Summary

We have reviewed a number of further topics in voting theory:

- *Weighted voting games:* power indices, the core, and their computational complexity
- *Proportional representation:* how to round quotas to calculate the number of seats a party is entitled to
- *Electronic voting:* how to make elections verifiable without compromising privacy
- Other topics: tournaments, algorithm design, approximation, parametrised complexity, repeated voting, ...

The Future

Voting is the central topic in computational social choice, but it is not the only one (we'll see some *other topics* over the next three weeks).

To keep track of *how the field develops* (and to find out about topics we did not cover), follow the COMSOC workshops and related events:

`http://www.illc.uva.nl/COMSOC/`

Some of the big *research challenges* for the coming years regarding voting theory in computational social choice (my personal view):

- Voting in combinatorial domains: languages and procedures.
- Develop a comprehensive theory for voting with ballots that need not be complete rankings of the full set of alternatives.
- Fully formalise (larger) parts of social choice theory and make them amenable to analysis via automated reasoning.
- (Maybe) integrate research in electronic voting and COMSOC.