Voting as Selection of the Most Representative Voter

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

[joint work with Umberto Grandi (Padova)]

Outline

- Examples
- Binary Aggregation with Integrity Constraints
- Representative-Voter Rules
- Approximation Results

Preference Aggregation

Expert 1: $\triangle \succ \bigcirc \succ \square$

Expert 2: $\bigcirc \succ \square \succ \triangle$

Expert 3: $\square \succ \triangle \succ \bigcirc$

Expert 4: $\square \succ \triangle \succ \bigcirc$

Expert 5: $\bigcirc \succ \square \succ \triangle$

?

Judgment Aggregation

	p	$p \to q$	q
Judge 1:	True	True	True
Judge 2:	True	False	False
Judge 3:	False	True	False

?

Multiple Referenda

	fund museum?	*fund school?*	*fund metro?*
Voter 1:	Yes	Yes	No
Voter 2:	Yes	No	Yes
Voter 3:	No	Yes	Yes

?

[Constraint: we have money for *at most two projects*]

General Perspective

The last example is actually pretty general. We can rephrase many aggregation problems as problems of *binary aggregation:*

Do you rank option △ above option ○?	Yes/No
Do you believe formula "$p \rightarrow q$" is true?	Yes/No
Do you want the new school to get funded?	Yes/No

Each problem domain comes with its own *rationality constraints:*

Rankings should be transitive and not have any cycles.

The accepted set of formulas should be logically consistent.

We should fund at most two projects.

The *paradoxes* we have seen show that the *majority rule* does not *lift* our rationality constraints from the *individual* to the *collective* level.

Binary Aggregation with Integrity Constraints

Basic terminology and notation:

- Set of *individuals* $\mathcal{N} = \{1, \ldots, n\}$; set of *issues* $\mathcal{I} = \{1, \ldots, m\}$.
- Corresponding set of *propositional symbols* $PS = \{p_1, \ldots, p_m\}$ and *propositional language* $\mathcal{L}_{PS}$ interpreted on $\mathcal{D} = \{0, 1\}^m$.
- An *aggregation rule* is a function $F : \mathcal{D}^n \to \mathcal{D}$. That is, each individual $i \in \mathcal{N}$ votes by submitting a *ballot* $B_i \in \mathcal{D}$.
- An *integrity constraint* is a formula $\mathrm{IC} \in \mathcal{L}_{PS}$ encoding a "rationality assumption". Ballot $B \in \mathcal{D}$ is *rational* iff $B \models \mathrm{IC}$.

Example

Our multiple-referenda example is formalised as follows:

- Three individuals: $\mathcal{N} = \{1, 2, 3\}$
- Three issues/prop. symbols: $\mathcal{I} = \{\texttt{museum}, \texttt{school}, \texttt{metro}\}$.
- Integrity constraint: $\mathrm{IC} = \neg(\texttt{museum} \wedge \texttt{school} \wedge \texttt{metro})$
- Profile: $\boldsymbol{B} = (B_1, B_2, B_3)$ with

$$\begin{aligned} B_1 &= (1, 1, 0) \\ B_2 &= (1, 0, 1) \\ B_3 &= (0, 1, 1) \end{aligned}$$

  Note that $B_i \models \mathrm{IC}$ for all $i \in \{1, 2, 3\}$

- However, $\mathrm{Maj}(\boldsymbol{B}) = (1, 1, 1)$ and $(1, 1, 1) \not\models \mathrm{IC}$.

Designing Good Aggregation Rules

We want to identify good methods for binary aggregation.

- Problem: the simple methods people use ("issue-wise majority") can lead to paradoxical outcomes.
- Problem: more sophisticated methods ("distance-based") are computationally intractable (as we will see).
- New idea: use an aggregation rule that identifies the "most representative" voter and just copies that voter's ballot.

Take-home message will be: simple, but works surprisingly well.

Distance-based Aggregation

How to avoid paradoxes?

$\rightarrow$ Only consider outcomes that respect the integrity constraint.
$\rightarrow$ Which one to pick?—the one "closest" to the individual inputs.

These considerations suggest the following rule:

- The (Hamming) *distance* between an individual input and the outcome is the number of "point decisions" on which they differ.
- Elect the (consistent/rational) outcome that *minimises* the sum of distances to the individual inputs! (+ break ties if needed)

For preference aggregation (with "point decisions" being pairwise rankings), this is the famous *Kemeny rule*. No rule is perfect, but many consider this one to be pretty good.

But: this is Θ_2^p-*complete* ("complete for parallel access to NP"). ☹

Taming the Complexity

Where does this complexity come from?
$\rightarrow$ We need to search through all candidate outcomes.

- there might be exponentially many of those
- for each of them, checking consistency might be nontrivial

An idea:

- restrict set of choices to a small set of candidate outcomes
- make sure you can be certain all candidate outcomes are consistent

The easiest way of doing this:

candidate outcomes = choices made by individuals ("support")

Example

Find the outcome that minimises the sum of distances for this profile:

Issue:	**1**	**2**	**3**
20 voters:	0	1	1
10 voters:	1	0	1
11 voters:	1	1	0

Solution: $(1, 1, 1)$. The distance is 41 (41 voters $\times$ 1 disagreement).

Note: same as majority outcome (as there's no integrity constraint).

Now suppose there's an IC that says that $(1, 1, 1)$ is not ok.

Example (continued)

Find the outcome that minimises the sum of distances for this profile:

Issue:	**1**	**2**	**3**
20 voters:	0	1	1
10 voters:	1	0	1
11 voters:	1	1	0

"Average voter" says: $(0, 1, 1)$.
The distance is 42 (20 with no disagreements + 21 with 2 each).

So: not much worse (42 *vs.* 41), but easier to find (choose from 3 rather than $2^3 = 8$ outcomes; all 3 known to be consistent *a priori*)

Additional Notation and Terminology

- *Hamming distance* between ballots: $H(B, B') = |\{j \in \mathcal{I} \mid b_j \neq b'_j\}|$ and between a ballot and a profile: $\mathcal{H}(B, \boldsymbol{B}) = \sum_{i \in \mathcal{N}} H(B, B_i)$.
- *Support* of profile $\boldsymbol{B}$: $\textsc{Supp}(\boldsymbol{B}) = \{B_1, \ldots, B_n\}$.

Rules Based on Representative Voters

Idea: Choose an outcome by first choosing a voter (based on the input profile) and then copying that voter's ballot.

$$\text{Fix } g : \mathcal{D}^n \to \mathcal{N}. \text{ Then let } F : \boldsymbol{B} \mapsto B_{g(\boldsymbol{B})}.$$

Good properties (of all these rules):

- *No paradoxes* ever, whatever the IC (not true for any other rule)
- *Unanimity* guaranteed [obvious]
- *Neutrality* guaranteed [maybe less obvious]
- *Low complexity* for natural choices of g

But:

- Includes some really bad rules, such as Arrovian *dictatorships:*

$$g \equiv i, \text{ i.e., } F : (B_1, \ldots, B_n) \mapsto B_i \text{ with } i \text{ being the dictator}$$

Two Representative-Voter Rules

The *average-voter rule* selects those individual ballots that minimise the Hamming distance to the profile:

$$\mathrm{AVR}(\boldsymbol{B}) \quad = \quad \operatorname*{argmin}_{B \in \textsc{Supp}(\boldsymbol{B})} \; \mathcal{H}(B, \boldsymbol{B})$$

<u>Remark:</u> if you replace the set $\textsc{Supp}(\boldsymbol{B})$ by $\mathrm{Mod}(\mathrm{IC})$, the set of *all* consistent outcomes, you obtain the full distance-based rule.

The *majority-voter rule* selects those individual ballots that minimise the Hamming distance to one of the majority outcomes:

$$\mathrm{MVR}(\boldsymbol{B}) \quad = \quad \operatorname*{argmin}_{B \in \textsc{Supp}(\boldsymbol{B})} \; \min\{H(B, B') \mid B' \in \mathrm{Maj}(\boldsymbol{B})\}$$

<u>Connections:</u>

- AVR related to *Kemeny* rule in voting/preference aggregation.
- MVR related to *Slater* rule in voting/preference aggregation.

Example

The AVR and the MVR really can give different outcomes:

Issue:	**1**	**2**	**3**	**4**	**5**	**6**
1 voter:	1	0	0	0	0	0
10 voters:	0	1	1	0	0	0
10 voters:	0	0	0	1	1	1
Maj:	0	0	0	0	0	0
MVR:	1	0	0	0	0	0
AVR:	0	1	1	0	0	0

Two More Representative-Voter Rules

We can also adapt *Tideman*'s ranked-pairs rule from voting theory. The *ranked-voter rule* (RVR) works as follows:

- order the issues by majority strength
- lock in issues in order of majority strength,
  whilst ensuring that the outcome remains within the support

The *plurality-voter rule* (PVR) selects the ballot chosen most often:

$$\text{PVR}(\boldsymbol{B}) \quad = \quad \underset{B \in \text{Supp}(\boldsymbol{B})}{\text{argmax}} \; |\{i \in \mathcal{N} \mid B = B_i\}|$$

The preference aggregation version of this rule has recently been proposed as a good maximum likelihood estimator by Caragiannis, Procaccia, and Shah ("modal ranking rule").

Approximation

F is said to be an *α-approximation* of F' if for every profile $\boldsymbol{B}$:

$$\max \mathcal{H}(F(\boldsymbol{B}), \boldsymbol{B}) \leqslant \alpha \cdot \min \mathcal{H}(F'(\boldsymbol{B}), \boldsymbol{B})$$

How well do our rules F approximate the *distance-based rule* F'?

- AVR: average-voter rule
- MVR: majority-voter rule
- RVR: ranked-voter rule
- PVR: plurality-voter rule
- Arrovian dictatorships $F_i : \boldsymbol{B} \mapsto B_i$

Good would be: α is a (small) *constant*

Bad would be: α depends on n or m, not bounded by any constant

Majority and Distance-based Rule

If the integrity constraint is "empty" ($\mathrm{IC} = \top$), the distance-based rule and the majority rule coincide:

$$\mathrm{DBR}^{\top} \quad = \quad \mathrm{Maj}$$

This is the hardest instance of the distance-based rule to approximate:

Lemma 1 *If* IC *logically entails* IC' ($\mathrm{IC} \models \mathrm{IC}'$), *then it is the case that* $\mathcal{H}(\mathrm{DBR}^{\mathrm{IC}}(\boldsymbol{B}), \boldsymbol{B}) \geqslant \mathcal{H}(\mathrm{DBR}^{\mathrm{IC}'}(\boldsymbol{B}), \boldsymbol{B})$ *for every profile* $\boldsymbol{B}$.

Proof: Both rules try to minimise the distance to the profile, but $\mathrm{DBR}^{\mathrm{IC}}$ has to choose from a smaller set of rational ballots. ✓

▶ For the rest of the talk, we will thus focus on approximating Maj (but exploring what additional mileage we can get out of stonger IC's is very interesting and not yet fully explored).

Very bad: Dictatorships

What's the worst possible scenario?

- one voter says $111 \cdots 111$, all others $(n-1)$ say $000 \cdots 000$
- majority rule would pick $000 \cdots 000$: *distance m*
- your rule picks $111 \cdots 111$: distance $m \cdot (n-1)$

Thus: worst approx. ratio for any rep-voter rule is $\frac{m \cdot (n-1)}{m} \in O(n)$

Arrovian dictatorships are maximally bad (unsurprinsngly):

Proposition 2 *Every Arrovian dictatorship $F_i : \boldsymbol{B} \mapsto B_i$ is a $\Theta(n)$-approximation of the majority rule.*

Proof: See above example, with dictator saying $111 \cdots 111$. ✓

Just as bad (!): RVR and PVR

Recall two of our more sophisticated rules:

- RVR: fix issues by majority strength, staying within support
- PVR: return most frequent ballot

Bad news:

Theorem 3 *RVR and PVR are $\Theta(n)$-approximations of* Maj.

Proof idea:

$$
\begin{array}{ll}
 & \overbrace{\quad\quad\quad\quad\quad\quad}^{n-2}\overbrace{\quad\quad\quad\quad\quad}^{m-(n-2)} \\
\text{Voter 1:} & 0\,1\,1\,1\,1 \cdots 1\,1\,1 \cdots\cdots\cdots 1 \\
\text{Voter 2:} & 1\,0\,1\,1\,1 \cdots 1\,1\,1 \cdots\cdots\cdots 1 \\
\vdots & \quad\quad\vdots \quad\quad\quad\quad \vdots \\
\text{Voter } n-2\text{:} & 1\,1\,1\,1\,1 \cdots 1\,0\,1 \cdots\cdots\cdots 1 \\
\text{Voter } n-1\text{:} & 1\,1\,1\,1\,1 \cdots 1\,1\,0 \cdots\cdots\cdots 0 \\
\text{Voter } n\text{:} & 1\,1\,1\,1\,1 \cdots 1\,1\,0 \cdots\cdots\cdots 0
\end{array}
$$

Good: MVR

Recall: the MVR selects the ballot closest to the majority outcome.

Theorem 4 *The* MVR *is a (strict)* 2*-approximations of* Maj.

Proof sketch: To simplify presentation, suppose there is only a single majority winner. W.l.o.g., suppose it is $(0, \ldots, 0)$.

Let m_i be the number of issues labelled as 1 by individual i. Let $i^\star$ be the voter selected by the MVR, i.e., $m_{i^\star} \leqslant m_i$ for all $i \in \mathcal{N}$.

If $m_{i^\star} = 0$, then we are done (approx. ratio 1). So suppose $m_{i^\star} \neq 0$.

We need to show:

$$\sum_{i \in \mathcal{N}} H(B_{i^\star}, B_i) \quad < \quad 2 \cdot \sum_{i \in \mathcal{N}} m_i$$

But this is the case:

- $H(B_{i^\star}, B_i) \leqslant m_{i^\star} + m_i \leqslant 2 \cdot m_i$ for all $i \neq i^\star$ (triangle inequality)
- $H(B_{i^\star}, B_{i^\star}) = 0 < 2 \cdot m_{i^\star}$ ✓

The best: AVR

Recall: the AVR selects the ballot closest to the input profile.

Thus, by definition:

Lemma 5 *The AVR approximimates Maj at least as well as any other representative-voter rule (thus: also a strict 2-approximation).*

Our most positive result:

Theorem 6 *Suppose m (the number of issues) is constant. Then the AVR is a $2\frac{m-1}{m}$-approximation of Maj.*

The same is *not* true for the MVR.

Recall that we can get better approximation ratios for $\mathrm{IC} \neq \top$.

Other Criteria for Comparison

Algorithmic *complexity:*

- Winner determination for the MVR is in $O(mn)$.
- Winner determination for the AVR is in $O(mn \log n)$.

F satisfies the axiom of *reinforcement* if

$$\textsc{Supp}(\boldsymbol{B}) = \textsc{Supp}(\boldsymbol{B}') \;\text{ and }\; F(\boldsymbol{B}) \cap F(\boldsymbol{B}') \neq \emptyset \quad \Rightarrow$$
$$F(\boldsymbol{B} \oplus \boldsymbol{B}') = F(\boldsymbol{B}) \cap F(\boldsymbol{B}')$$

Theorem 7 *The* AVR *satisfies reinforcement, but the* MVR *does not.*

Last Slide

This work is part of a larger effort to better understand the powerful framework of *binary aggregation with integrity constraints*. The focus today has been on identifying good and simple rules to use in practice.

- Simple (maybe simplistic) idea: pick a representative voter + copy
- Surprisingly, this can work very well; we can get good properties:
  - guarantee to never encounter a paradox
  - low complexity
  - good social choice-theoretic axioms (though not independence)
  - for some: good approximation ratios wrt. distance-based rule

U. Endriss and U. Grandi. Binary Aggregation by Selection of the Most Representative Voter. *Proc. AAAI-2014*.