Binary Aggregation by Selection of the Most Representative Voter

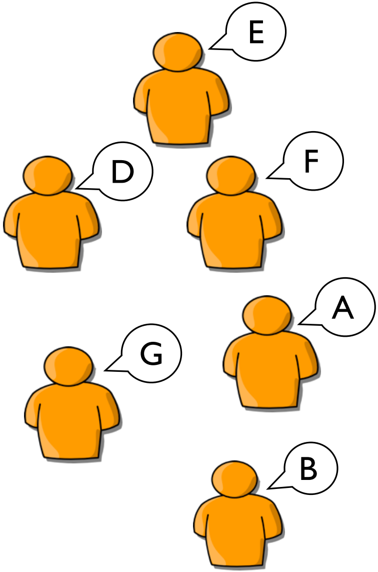
Umberto Grandi

Department of Mathematics
University of Padova

14 May 2014

[Joint work with Ulle Endriss]

Selection of the Closest Opinion



$$\underset{\{o_i | i \in \mathcal{N}\}}{\operatorname{argmin}} \; d(o_i, o_1, \ldots, o_n)$$

Opinion A

Opinion B

Opinion C

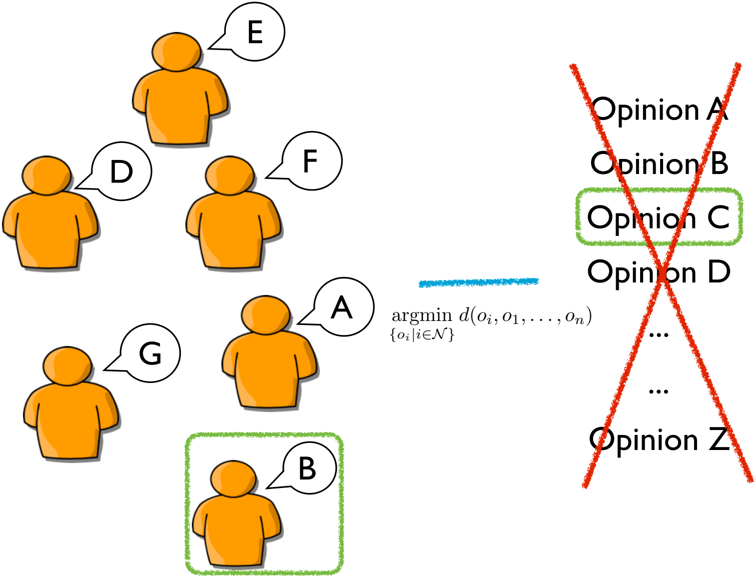
Opinion D

...

...

Opinion Z

Selection of the Most Representative Voter



Outline

1. A general framework for aggregation problems:
   - Binary aggregation with integrity constraints
   - Preferences, judgments, multi-issue elections...
   - Generalised dictatorships
2. Selection of the most representative voter:
   - Average voter rule (AVR)
   - Majority voter rule (MVR)
   - Ranked voter rule (RVR)
3. Properties of most-representative-voter rules:
   - Approximation results
   - Computational complexity
   - Axiomatic properties

Binary Aggregation

Ingredients:

- A finite set $\mathcal{N}$ of individuals
- A finite set $\mathcal{I} = \{1, \ldots, m\}$ of issues
- A boolean combinatorial domain: $\mathcal{D} = \{0, 1\}^{\mathcal{I}}$

Definition

An aggregation procedure is a function $F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}$ mapping each profile of ballots $\boldsymbol{B} = (B_1, \ldots, B_n)$ to an element of the domain $\mathcal{D}$.

Wilson (1975), Dokow and Holzman (JET 2010), Grandi and Endriss (AIJ 2013)

Integrity Constraints

A propositional language $\mathcal{L}$ to define the subset of rational ballots in $\{0,1\}^{\mathcal{I}}$:

- One propositional symbol for every issue: $PS = \{p_1, \ldots, p_m\}$
- $\mathcal{L}_{PS}$ closed under connectives $\wedge$, $\vee$,$\neg$, $\rightarrow$ the set of atoms PS

Given an integrity constraint IC $\in \mathcal{L}_{PS}$, a rational ballot is $B \in \mathrm{Mod}(\mathrm{IC})$

Integrity Constraints

A propositional language $\mathcal{L}$ to define the subset of rational ballots in $\{0,1\}^{\mathcal{I}}$:

- One propositional symbol for every issue: $PS = \{p_1, \ldots, p_m\}$
- $\mathcal{L}_{PS}$ closed under connectives $\wedge$, $\vee$,$\neg$, $\rightarrow$ the set of atoms PS

Given an integrity constraint IC $\in \mathcal{L}_{PS}$, a rational ballot is $B \in \mathrm{Mod}(\mathrm{IC})$

Example: Three agents with sensors

Perform action A if both parameters T_1 and T_2 exceed the thresholds.
Propositional constraint: $\mathrm{IC} = (p_{T_1} \wedge p_{T_2}) \rightarrow p_A$

Individual 1 submits $B_1 = (1,1,1)$:	B_1 satisfies IC ✓
Individual 2 submits $B_2 = (0,1,0)$:	$B_2 \models$ IC ✓
Individual 3 submits $B_3 = (1,0,0)$:	$B_3 \models$ IC ✓

Majority aggregation outputs $(1,1,0)$: IC not satisfied.

Preference Aggregation as Binary Aggregation

Agent 1	$A > B > C$
Agent 2	$B > C > A$
Agent 3	$C > A > B$
Maj	$A > B > C > A$!!

Condorcet
Paradox (1785)

Preferences as
binary ballots
+ integrity constraint

	$A > B$	$B > C$	$A > C$
Agent 1	1	1	1
Agent 2	0	1	0
Agent 3	1	0	0
Maj	1	1	0

The Discursive Dilemma

Agent 1	$\{\alpha, \beta, \alpha \wedge \beta\}$
Agent 2	$\{\neg\alpha, \beta, \neg(\alpha \wedge \beta)\}$
Agent 3	$\{\alpha, \neg\beta, \neg(\alpha \wedge \beta)\}$
Maj	$\{\alpha, \beta, \neg(\alpha \wedge \beta)\}$

Doctrinal Paradox

Kornhauser and Sager (1986)

Judgments as binary ballots + integrity constraint

$$\text{IC} = \neg(p_\alpha \wedge p_\beta \wedge p_{\neg(\alpha\wedge\beta)})$$

	p_α	p_β	$p_{\alpha\wedge\beta}$
Agent 1	1	1	1
Agent 2	0	1	0
Agent 3	1	0	0
Maj	1	1	0

Avoid paradoxes? Characterisation results and generalised dictatorship

Proposition - Majority rule

The majority rule does not generate a paradox with respect to IC if and only if IC is equivalent to a conjunction of clauses of size $\leqslant 2$ *(i.e.,* 2-CNF*)*

Avoid paradoxes? Characterisation results and generalised dictatorship

Proposition - Majority rule

The majority rule does not generate a paradox with respect to IC if and only if IC is equivalent to a conjunction of clauses of size $\leqslant 2$ *(i.e.,* 2-CNF*)*

How to avoid all paradoxes?

Proposition - Avoid all paradoxes

An aggregation procedure does not generate a paradox with respect to all IC if and only if it copies the ballot of a (possibly different) individual in every profile.

Grandi and Endriss, Lifting Integrity Constraints in Binary Aggregation. *AIJ*, 2013.

Distance-based rules in judgment aggregation

Definition - Distance-based rule

The DBR *(aka Kemeny rule, Prototype) picks the consistent ballots minimising the sum of the Hamming distances to the individual ballots.* Θ_2^p*-complete*

Definition - Slater rule

The Slater rule (aka Endpoint) picks the consistent ballots minimising the Hamming distance to the outcome of the majority rule. NP-hard (at least)

Definition - Ranked agenda

The ranked-agenda rule picks the consistent ballots obtained by sequential majority following the order given by the strength of acceptance. Δ_2^p*-hard.*

Endriss, Grandi and Porello. Complexity of Judgment Aggregation, *JAIR*, 2012.
Brandt et Al. The Computational Complexity of Choice Sets. *Mathematical Logic Quarterly*, 2009.
Lang and Slavkovijk, *ECAI-2014*.

Selection of the Most Representative Voter

Restrict the search space to $\textsc{Supp}(\boldsymbol{B}) = \{B_1, \ldots, B_n\}$

Definition

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

$$\mathrm{AVR}(\boldsymbol{B}) = \operatorname*{argmin}_{B \in \textsc{Supp}(\boldsymbol{B})} \sum_{i \in \mathcal{N}} H(B, B_i)$$

Definition

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

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

The RVR is defined in a similar way...

An Example

The AVR, the MVR and the majority rule can give radically different results:

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

Hamming distance of AVR from the profile: 53
Hamming distance of MVR from the profile: 70

Observation

$\mathcal{H}(\mathrm{AVR}(\boldsymbol{B}), \boldsymbol{B}) \leqslant \mathcal{H}(\mathrm{MVR}(\boldsymbol{B}), \boldsymbol{B})$ *where* $\mathcal{H}(B, \boldsymbol{B}) = \sum_i H(B, B_i)$

Approximation of an ideal rule

Can we compare the outcome of the ideal rule – DBR aka Kemeny – with that of our AVR, MVR, RVR?

Definition - Approximation

F is said to be an α-approximation of $\mathrm{DBR}^{\mathrm{IC}}$ if for every profile $\boldsymbol{B}$:

$$\mathcal{H}(F(\boldsymbol{B}), \boldsymbol{B}) \leqslant \alpha \cdot \mathcal{H}(\mathrm{DBR}^{\mathrm{IC}}(\boldsymbol{B}), \boldsymbol{B})$$

Good approximation if α is a constant.

Preliminary facts about the distance-based procedure

The definition of the distance-based rule depends on the constraint:

$$\mathrm{DBR}^{\mathrm{IC}}(\boldsymbol{B}) = \underset{B \in \mathrm{Mod}(\mathsf{IC})}{\operatorname{argmin}} \sum_{i \in \mathcal{N}} H(B, B_i)$$

In particular, $\mathrm{DBR}^{\top} = \mathrm{Maj}$ (the majority rule). With stronger constraints?

Lemma

If IC *entails* IC', *then* $\mathcal{H}(\mathrm{DBR}^{\mathrm{IC}}(\boldsymbol{B}), \boldsymbol{B}) \geqslant \mathcal{H}(\mathrm{DBR}^{\mathrm{IC}'}(\boldsymbol{B}), \boldsymbol{B})$ *for every profile* $\boldsymbol{B} \in \mathrm{Mod}(\mathrm{IC})^n$.

And a baseline result:

Proposition

Every rule based on the most representative voter is an $O(n)$*-approximation of the* $\mathrm{DBR}^{\mathrm{IC}}$.

Negative result

Recall that n is the number of individuals, m is the number of issues.

Theorem

The RVR *is a* $\Theta(n)$*-approximation of* Maj *(even if* m *is bounded).*

Proof. The upper bound is given by the result on the previous slide. The lower bound is obtained by showing a family of profiles where the result of the RVR is n-far from that of Maj.

Positive Results

Theorem

The AVR *and the* MVR *are strict 2-approximations of the* $\mathrm{DBR}^{\mathrm{IC}}$.

The AVR can go closer than 2 if m is bounded or the IC is restrictive:

Theorem

Let m *be constant. Then the* AVR *is an* α*-approximation of the* $\mathrm{DBR}^{\mathrm{IC}}$ *with* $\alpha = 2\frac{m-1}{m}$ *for any integrity constraint* IC.

Theorem

Let m *be constant and let* IC *be a conjunction of* k *distinct literals. Then the* AVR *is an* α*-approximation of the* $\mathrm{DBR}^{\mathrm{IC}}$ *with* $\alpha = 2\frac{m-k-1}{m-k}$.

The MVR cannot do better, even with a bounded number of issues.

Computational Complexity

Recall that m is the number of issues; n is the number of voters.

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

Winner determination for the MVR is in $O(mn)$

Conclusion? Both rules are easy to compute (MVR is easier)

Axiomatic Properties

Rules based on the most representative voter satisfy interesting properties:

- No paradox ever, whatever the IC (no other rule has this property)
- Unanimity guaranteed (obvious)
- Neutrality guaranteed (less obvious)

F satisfies reinforcement if for any two profiles $\boldsymbol{B}$ and $\boldsymbol{B}'$ such that:

- $\textsc{Supp}(\boldsymbol{B}) = \textsc{Supp}(\boldsymbol{B}')$
- $F(\boldsymbol{B}) \cap F(\boldsymbol{B}') \neq \emptyset$

we have that $F(\boldsymbol{B} \oplus \boldsymbol{B}') = F(\boldsymbol{B}) \cap F(\boldsymbol{B}')$

Theorem

The AVR satisfies reinforcement, but the MVR does not.

Conclusions

Characterisation results in binary aggregation with integrity constraints suggests novel simple procedures to be used in practice:

- AVR: voters minimizing average distance
- MVR: voters minimizing distance from majority
- RVR: voters chosen by sequential majority following strength order

Very good properties (AVR and MVR):

- Outcome will never be paradoxical
- Very low complexity
- Social-choice theoretic properties (not independence!)
- 2-approximation of distance-based rule (aka Kemeny)

Not all definitions work well: RVR is a bad approximation.

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