Voting on combinatorial domains

Jérôme Lang
LAMSADE, CNRS – Université Paris-Dauphine

FET-11, session on Computational Social Choice

A key question: *structure* of the set $\mathcal{X}$ of candidates?

Example 1 choosing a common menu:

$$\begin{aligned} \mathcal{X} = \quad & \{\texttt{asparagus risotto, foie gras}\} \\ \times \quad & \{\texttt{roasted chicken, vegetable curry}\} \\ \times \quad & \{\texttt{white wine, red wine}\} \end{aligned}$$

Example 2 *multiple referendum*: a local community has to decide on several interrelated issues (should we build a swimming pool or not? should we build a tennis court or not?)

Example 3 *choosing a joint plan*. A group of friends has to travel together to a sequence of possible locations, given some constraints on the possible sequences.

Example 4 *committee election*; choose three representatives out of 6 candidates.

$$\mathcal{X} = \{A \mid A \subseteq \{a, b, c, d, e, f\},\ |A| \leq 3\}$$

Example 1 *common menu*

$$\begin{aligned} \mathcal{X} = & \quad \{\texttt{asparagus risotto, foie gras}\} \\ & \times \{\texttt{roasted chicken, vegetable curry}\} \\ & \times \{\texttt{white wine, red wine}\} \end{aligned}$$

Example 2 *multiple referendum*

$$\mathcal{X} = \{\texttt{swimming pool, no swimming pool}\} \times \{\texttt{tennis, no tennis}\}$$

Example 3 *joint plan / group traveling*

$$\mathcal{X} = \text{set of all possible allowed paths in the graph}$$

Example 4 *committee election*

$$\mathcal{X} = \{A \mid A \subseteq \{a,b,c,d,e,f\},\ |A| \leq 3\}$$

Examples 1-4: *voting on a combinatorial domain*.

Set of alternatives: $\mathcal{X} = D_1 \times ... \times D_p$ where

- $\mathcal{V} = \{X_1, \ldots, X_p\}$ set of *variables*, or *issues*;
- D_i is a finite value domain for variable X_i)

How should such a vote be conducted?

How should such a vote be conducted?

1. **don’t bother and vote simultaneously on each variable.**

How should such a vote be conducted?

1. **don't bother and vote simultaneously on each variable**

Example

2 binary variables S (build a new swimming pool), T (build a new tennis court)

voters 1 and 2	$S\bar{T} \succ \bar{S}T \succ \bar{S}\bar{T} \succ ST$
voters 3 and 4	$\bar{S}T \succ S\bar{T} \succ \bar{S}\bar{T} \succ ST$
voter 5	$ST \succ S\bar{T} \succ \bar{S}T \succ \bar{S}\bar{T}$

How should such a vote be conducted?

1. **don't bother and vote simultaneously on each variable**.

Example

2 binary variables S (build a new swimming pool), T (build a new tennis court)

voters 1 and 2	$S\bar{T} \succ \bar{S}T \succ \bar{S}\bar{T} \succ ST$
voters 3 and 4	$\bar{S}T \succ S\bar{T} \succ \bar{S}\bar{T} \succ ST$
voter 5	$ST \succ S\bar{T} \succ \bar{S}T \succ \bar{S}\bar{T}$

Problem 1: voters 1-4 feel ill at ease reporting a preference on $\{S, \bar{S}\}$ and $\{T, \bar{T}\}$

How should such a vote be conducted?

1. **don't bother and vote simultaneously on each variable**.

Example

2 binary variables S (build a new swimming pool), T (build a new tennis court)

voters 1 and 2	$S\bar{T} \succ \bar{S}T \succ \bar{S}\bar{T} \succ ST$
voters 3 and 4	$\bar{S}T \succ S\bar{T} \succ \bar{S}\bar{T} \succ ST$
voter 5	$ST \succ S\bar{T} \succ \bar{S}T \succ \bar{S}\bar{T}$

Problem 1: voters 1-4 feel ill at ease reporting a preference on $\{S, \bar{S}\}$ and $\{T, \bar{T}\}$

Problem 2: suppose they do so by an "optimistic" projection

- voters 1, 2 and 5: S; voters 3 and 4: $\bar{S} \Rightarrow$ decision = S;
- voters 3,4 and 5: T; voters 1 and 2: $\bar{T} \Rightarrow$ decision = T.

Alternative ST is chosen although it is the worst alternative for all but one voter.

Multiple election paradoxes arise as soon as some voters have nonseparable preferences

How should such a vote be conducted?

1. don't bother and vote simultaneously on each variable.
2. **ask voters to specify their preference relation by ranking all alternatives *explicitly***.

$\mathcal{V} = \{X_1, \ldots, X_p\}$; $\mathcal{X} = D_1 \times \ldots \times D_p$

There are $\Pi_{1 \leq i \leq p} |D_i|$ alternatives.

Example: in a committee election with 15 candidates, there are $2^{10} = 32768$ alternatives.

As soon as there are more than three or four variables, explicit preference elicitation is irrealistic.

How should such a vote be conducted?

1. don't bother and vote simlutaneously on each variable.
2. ask voters to specify their preference relation by ranking all alternatives *explicitly*.
3. **ask voters to report only a small part of their preference relation and appply a voting rule that needs this information only, such as plurality.**

5 voters, 2^6 alternatives; rule : plurality

001010: 1 vote; 010111: 1 vote; 011000: 1 vote; 101001: 1 vote; 111000: 1 vote
all other candidates : 0 vote.

Results are generally completely nonsignificant as soon as the number of alternatives is much higher than the number of voters ($2^p \gg n$).

How should such a vote be conducted?

1. don't bother and vote simultaneously on each variable.
2. ask voters to specify their preference relation by ranking all alternatives *explicitly*.
3. ask voters to report only a small part of their preference relation and appply a voting rule that needs this information only, such as plurality.
4. **ask voters their preferred alternative(s) and complete them automatically using a predefined *distance*.**

- the agent specifies only her preferred alternative $\vec{x}$
- and her preference is completed by $\vec{y} \succ \vec{z}$ if and only if $\vec{y}$ is closer to $\vec{x}$ than $\vec{z}$

Example: Hamming distance d_H

- $\vec{x} = ab\overline{c}$
- $ab\overline{c} \succ [abc \sim a\overline{b}\overline{c} \sim \overline{a}b\overline{c}] \succ [a\overline{b}c \sim \overline{a}\overline{b}\overline{c} \sim \overline{a}bc] \succ \overline{a}\overline{b}c$

Needs an important domain restriction + can be computationally difficult

How should such a vote be conducted?

1. don't bother and vote simultaneously on each variable.
2. ask voters to specify their preference relation by ranking all alternatives *explicitly*.
3. ask voters to report only a small part of their preference relation and appply a voting rule that needs this information only, such as plurality.
4. ask voters their preferred alternative(s) and complete them automatically using a predefined *distance*.
5. ***sequential voting*** **: decide on every variable one after the other, and broadcast the outcome for every variable before eliciting the votes on the next variable.**

Sequential voting

voters 1 and 2	$S\bar{T} \succ \bar{S}T \succ \bar{S}\bar{T} \succ ST$
voters 3 and 4	$\bar{S}T \succ S\bar{T} \succ \bar{S}\bar{T} \succ ST$
voter 5	$ST \succ S\bar{T} \succ \bar{S}T \succ \bar{S}\bar{T}$

Fix the order $S > T$.

Step 1 elicit preferences on $\{S, \bar{S}\}$

if voters report preferences optimistically: $3 : S \succ \bar{S}$; $2 : \bar{S} \succ S$

Step 2 compute the local outcome and broadcast the result

$$S$$

Step 3 elicit preferences on $\{T, \bar{T}\}$ given the outcome on $\{S, \bar{S}\}$

4: $S : \bar{T} \succ T$; 1: $S : T \succ \bar{T}$

Step 4 compute the final outcome

$$S\bar{T}$$

Sequential voting

- The outcome may depend on the order: the chair partially controls the process
- Much better than simultaneous voting but partially suffers from the same problems (voters may experience regret after the final outcome is known)

How should such a vote be conducted?

1. don’t bother and vote simlutaneously on each variable.
2. ask voters to specify their preference relation by ranking all alternatives *explicitly*.
3. ask voters to report only a small part of their preference relation and appply a voting rule that needs this information only, such as plurality.
4. ask voters their preferred alternative(s) and complete them automatically using a predefined *distance*.
5. *sequential voting* : decide on every variable one after the other, and broadcast the outcome for every variable before eliciting the votes on the next variable.
6. **use a *compact preference representation language* in which the voters’ preferences are represented in a concise way.**

   potentially expensive in elicitation and/or computation

How should such a vote be conducted?

Conclusions: we have to make trade-offs between:

- strong domain restrictions
- inefficiency
- high computational cost
- high communication cost

$\Rightarrow$ design "efficient" *elicitation protocols*; try to minimize the amount of communication between the voters and the central authority

$\Rightarrow$ develop sophisticated algorithms

$\Rightarrow$ identify restrictions under which the elicitation cost and/or the complexity cost are reasonable/