Elicitation of the parameters of a general ordinal sorting model in a conjoint measurement setting

Marc Pirlot and Eda Ersek

Faculté Polytechnique de Mons
Belgium

Catania Workshop COST IC0602
Algorithmic Decision Theory   April 16, 2008

Motivation

D. Bouyssou and myself (M.P.) (together with others like S. Greco, B. Matarazzo, R. Slowinski, Th. Marchant, . . .) have been trying to develop axiomatic foundations for outranking methods for many years.

And we have done so in a conjoint measurement framework.

Why? To imitate what had been done for the *additive value (utility) function* model ?

In a sense, the answer is “YES” but there are also good reasons . . .

Axiomatic analysis is the key to the elaboration of rigorous elicitation methods

The main virtues of an axiomatic analysis are that

- the model is completely described by one or several system(s) of properties
- the axioms may provide means for testing the applicability of the model
- it focuses the attention on primitives (e.g. tradeoffs, marginal preferences, ...) on which the elicitation procedure can rely

Illustration : using the *Standard sequence* method to build an additive value fonction

$$x \succsim y \Leftrightarrow \sum_{i=1}^{n} u_i(x_i) \geq \sum_{i=1}^{n} u_i(y_i)$$

Well-designed sequence of *indifference judgments*

Example (Bouyssou et al. [00]): a student buys a second-hand sportive car

A standard sequence is used for building an (approximate) value function on the criterion "acceleration" on the basis of a *standard* price difference: 15,000 € to 16,000 €

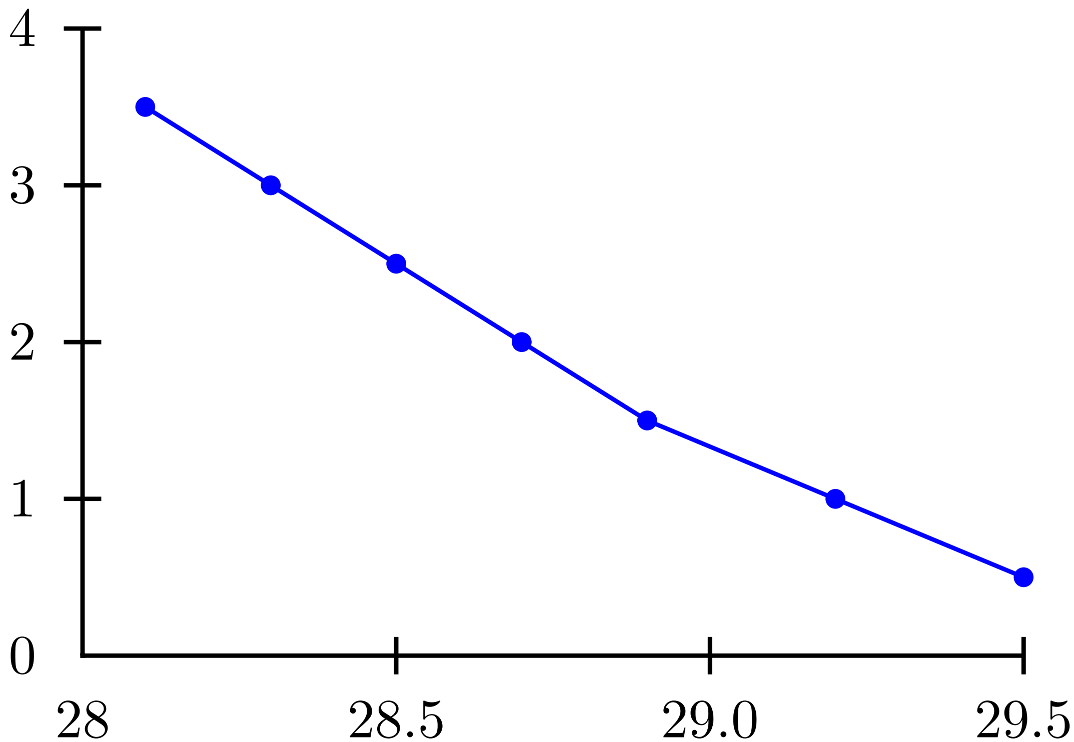
As a sportsman, the student is interested in descending under 29 sec.

Two attributes considered here: *price* and *acceleration* (time needed to cover 1 km, in seconds)

The DM is asked to equilibrate a balance in the following cases in turn:

- (15,000 ; 29.5) $\sim$ (16,000; ?) Answer = 29.2
- (15,000 ; 29.2) $\sim$ (16,000; ?) Answer = 28.9
- (15,000 ; 28.9) $\sim$ (16,000; ?) Answer = 28.7
- ...
- (15,000 ; 28.3) $\sim$ (16,000; ?) Answer = 28.1

This sequence allows to construct an approximation of the function u_2 recoding the attribute “acceleration”


4
3
2
1
0
28
28.5
29.0
29.5

Question : Is it possible to do something similar with outranking methods and especially with ELECTRE-TRI ?

We use a simplified and axiomatized version of ELECTRE-TRI: the Non-Compensatory Sorting Model

Summary

- The Non-Compensatory Sorting Model
- Elicitation issues
- Eliciting the set of satisfactory levels
- Eliciting the set of sufficient coalitions
- Further research and perspectives

The Non-Compensatory Sorting Model

What is NCSM ? Conjoint measurement model for sorting alternatives in pre-defined ordered categories

- inspired by Electre-Tri (pessimistic version): the assignment of an alternative to a category is done by comparing the alternative to a limiting profile by using an outranking rule
- in conjoint measurement models, alternatives are usually the elements of a cartesian product.
- here, the primitive object is an assignment of the alternatives in ordered categories, i.e. an ordered partition
- the model consists in specifying a particular way of making the assignment (= assignment rule = model)

- one seeks to describe completely this type of assignment by a set of characteristic properties (axioms)

- Developed by D. Bouyssou and Th. Marchant (EJOR 07, 2 papers: two categories & more than two categories)
- Following previous work by Goldstein (JMP 91) and Greco et al (01), Slowinski et al (Control & Cybernetics 02)

The setting

- The alternatives are all the elements of a product set $X = X_1 \times X_2 \times \ldots \times X_n$;
- interpretation: an alternative x is identified to the vector $(x_1, \ldots, x_n)$ of its evaluations on the set $N = \{1, \ldots, n\}$ of attributes
- One can thus build an alternative by mixing up other alternatives x and y; for instance:
  - for a subset $J \subseteq N$ of criteria, we may consider $z = (x_J, y_{-J})$
  - and, abusing notation, we shall often consider $z = (x_i, y_{-i})$

Assignment

- An assignment in two (ordered) categories is an ordered bipartition of X, $\langle \mathcal{A}, \mathcal{U} \rangle$ with:
  - $\mathcal{A}$ denoting the set of “acceptable” alternatives and
  - $\mathcal{U}$ denoting the set of “unacceptable” alternatives
- An assignment in r (ordered) categories is a partition of X in r ordered classes $\langle \mathcal{C}^1, \ldots \mathcal{C}^r \rangle$ with:
  - $\mathcal{C}^1$ denoting the “worst” category of alternatives and
  - $\mathcal{C}^r$ denoting the “best” category of alternatives

In the sequel, we focus on sorting in 2 categories. Sorting in more than 2 categories can be seen as repeated sorting in 2 categories

The case of sorting in two categories

Definition (Bouyssou, Marchant EJOR 07)
An ordered partitioning $\langle \mathcal{A}, \mathcal{U} \rangle$ of X has a representation in the *Non-Compensatory Sorting Model* if

- on each dimension i, there is a set $\mathcal{A}_i \subseteq X_i$ and
- there is a family $\mathcal{F}$ of subsets of N that is final, (i.e. $I \in \mathcal{F}$ and $J \supseteq I \;\Rightarrow\; J \in \mathcal{F}$)

such that:

$$x \in \mathcal{A} \quad \Leftrightarrow \quad \{i \in N : x_i \in \mathcal{A}_i\} \in \mathcal{F}$$

Interpretation

- $\mathcal{A}_i$ is the set of "satisfactory" levels on dimension i
- $\mathcal{F}$ is the family of "sufficient coalitions" of criteria

Justification for the term "Non-Compensatory":
Very rough distinction among "levels" on the scale of each criterion i: only two classes of equivalence

Remark:
We do not consider vetoes here (while Bouyssou-Marchant do also characterize the "Non-Compensatory Sorting Model with Veto")

Characterization result

Theorem (Bouyssou, Marchant EJOR 07)
An ordered partitioning $\langle \mathcal{A}, \mathcal{U} \rangle$ of X has a representation in the *Non-Compensatory Sorting Model* iff

- it is linear and
- it is $2-$graded

Interpretation:

- Linearity is equivalent to assuming that relation $\succsim_i$, defined below, is a complete preorder on each X_i:

$$x_i \succsim_i y_i \;\Leftrightarrow\; [\forall b_{-i},\; (y_i, b_{-i}) \in \mathcal{A} \;\Rightarrow\; (x_i, b_{-i}) \in \mathcal{A}]$$

- under linearity, $2-$gradedness is equivalent to assuming that $\succsim_i$ has at most two equivalence classes

Remark The axioms of linearity and $2-$gradedness are, in principle, testable

Relationship with ELECTRE-TRI

Consider a simple version of ELECTRE-TRI ("pessimistic version") in which

- there are only two classes
- the preference and the indifference thresholds are equal
- there are no discordance effects

For each i, there is a semiorder S_i on X_i ("at least as good" relation on X_i). A weight w_i is associated to each criterion i; these weights can be supposed to be normalized. Let λ be a number between 1/2 and 1.

Given a profile $p = (p_1, \ldots, p_n) \in X$, ELECTRE-TRI determines that

$$x \in \mathcal{A} \Leftrightarrow \sum_{i: x_i S_i p_i} w_i \geq \lambda$$

The obtained partition $\langle \mathcal{A}, \mathcal{U} \rangle$ is representable in the NCSM model with

- $\mathcal{A}_i = \{x_i : x_i S_i p_i\}$
- $\mathcal{F} = \{I \subseteq N : \sum_{i \in I} w_i \geq \lambda\}$

We also have $x_i \succ_i y_i$ iff $x_i S_i p_i$ and Not $[y_i S_i p_i]$

Elicitation of $\mathcal{A}_i$ and $\mathcal{F}$

Assumptions

- for all i, X_i is an interval (a_i, b_i) of the real line, with $a_i < b_i$
- the unknown partition $\langle \mathcal{A}, \mathcal{U} \rangle$ of $X = \prod X_i$ is representable in the NCSM (testable)
- the natural order $\geq_i$ on X_i is compatible with the weak order $\succsim_i$, i.e.

$$x_i \geq_i y_i \;\Rightarrow\; x_i \succsim_i y_i$$

Influent criteria

Definition A criterion i is influent if there are $x_i, y_i \in X_i$ and $z_{-i} \in X_{-i}$ such that

$$(x_i, z_{-i}) \in \mathcal{A} \text{ and } (y_i, z_{-i}) \notin \mathcal{A}$$

In such a case, $x_i \in \mathcal{A}_i$ and $y_i \notin \mathcal{A}_i$.

Assume that we know $\mathcal{F}$, the set of sufficient coalitions. We can find out which criteria are influent by looking at $\mathcal{F}$. Indeed, i is influent if there is $I \in \mathcal{F}$ such that

$$i \in I \text{ and } I \setminus \{i\} \notin \mathcal{F}$$

Note also that

$$(x_i, z_{-i}) \in \mathcal{A} \text{ and } (y_i, z_{-i}) \notin \mathcal{A}$$

entails

$$(b_i, z_{-i}) \in \mathcal{A} \text{ and } (a_i, z_{-i}) \notin \mathcal{A}$$

since $\geq_i$ on $X_i = (a_i, b_i)$ is compatible with $\succsim_i$ and

$$b_i \geq_i x_i >_i y_i \geq_i a_i$$

Eliciting $\mathcal{A}_i$ knowing $\mathcal{F}$

Assume that $\mathcal{F}$ has been determined

Remember that $X_j = (a_j, b_j)$, with $a_j < b_j$. For all $J \subseteq N$, let b_J denote a vector of evaluations along the criteria in J, in which all values are maximal (w.r.t. $\geq_i$), i.e. equal to b_j for all $j \in J$. And a_J will denote a vector of evaluations on J, in which all evaluations are minimal, i.e. equal to a_j for all $j \in J$.

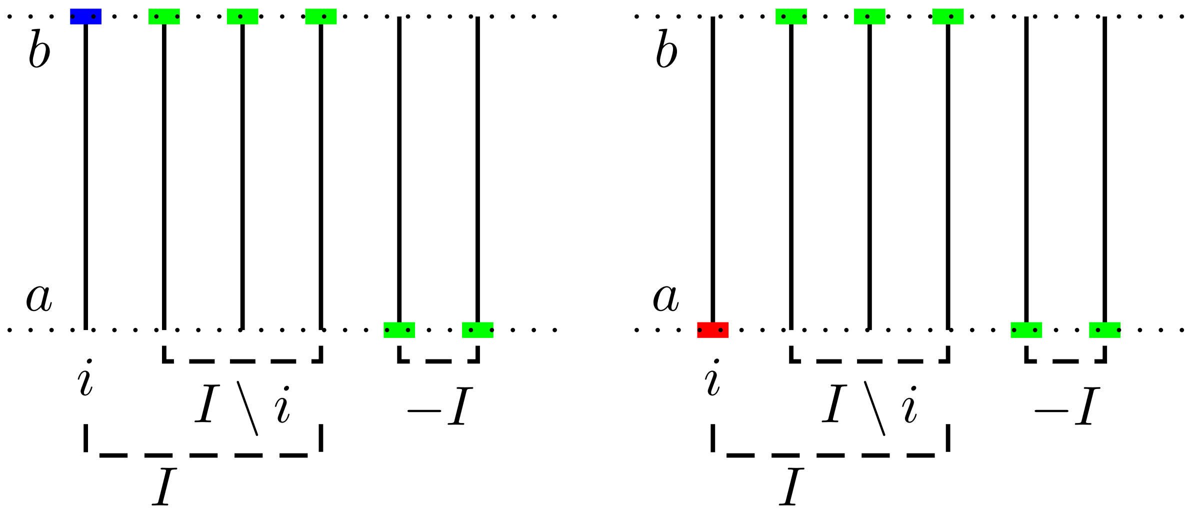
Let i be an influent criterion. Then there is $I \in \mathcal{F}$ such that
$i \in I$ and $I \setminus \{i\} \notin \mathcal{F}$

We have:

$$(b_I, a_{-I}) \in \mathcal{A} \text{ and } (a_i, b_{I \setminus i}, a_{-I}) \notin \mathcal{A}$$

Graphically,

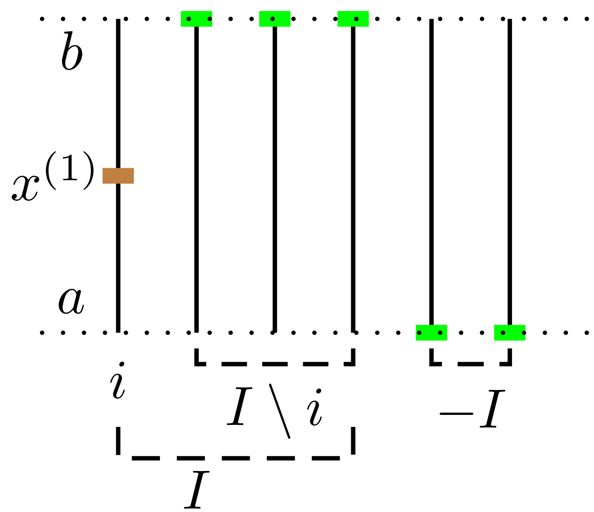
$$(b_I, a_{-I}) \in \mathcal{A} \quad \text{and} \quad (a_i, b_{I\setminus i}, a_{-I}) \notin \mathcal{A}$$



We want to determine (approximately) $\mathcal{A}_i = [z_i, b_{-i}]$ or $]z_i, b_{-i}]$, for some $z_i \in [a_i, b_i]$

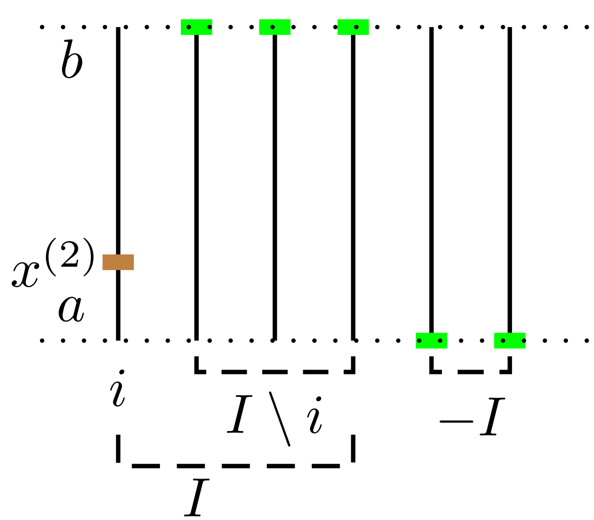
We engage into a dichotomic procedure: let $x_i^{(1)}$ be the middle of $[a_i, b_i]$

Ask whether

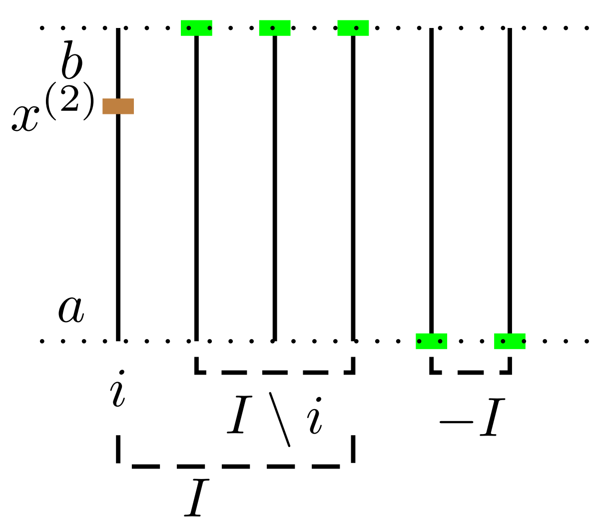


$\in \mathcal{A}$?

If YES $\longrightarrow$

Ask whether  $\in \mathcal{A}$?

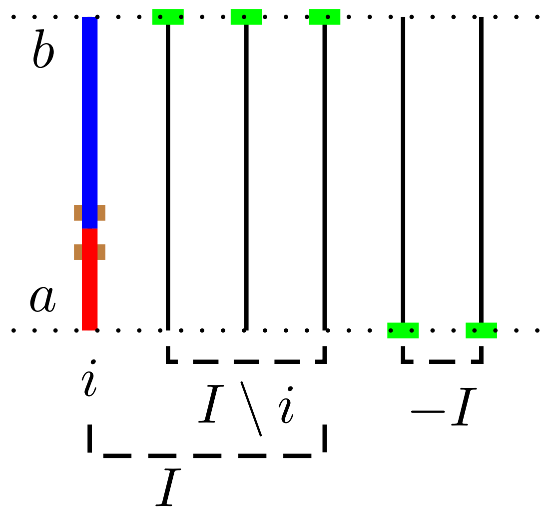
If NO $\longrightarrow$

Ask whether  $\in \mathcal{A}$?

Finally, when we have, after k steps:

- $(x_i^{(k)}, b_{I \setminus i}, a_{-I}) \in \mathcal{A}_i$ and
- $(x_i^{(k+1)}, b_{I \setminus i}, a_{-I}) \notin \mathcal{A}_i$

we know that $z_i \in]x_i^{(k+1)}, x_i^{(k)}]$



Precision: $(\frac{1}{2})^{k+1} \times (b_i - a_i)$

Remarks

- In case there are only finitely many alternatives to be sorted, it is possible to proceed with dichotomization up to the point of determining whether any *observed* level of X_i is in $\mathcal{A}_i$ or not
- if i is not an influent criterion, it makes no difference saying that $\mathcal{A}_i = X_i$ or $\mathcal{A}_i = \emptyset$

Eliciting $\mathcal{F}$

In Electre-Tri, it is assumed that being a sufficient coalition can be determined by using (normalized) weights w_i and a threshold λ in the following rule:

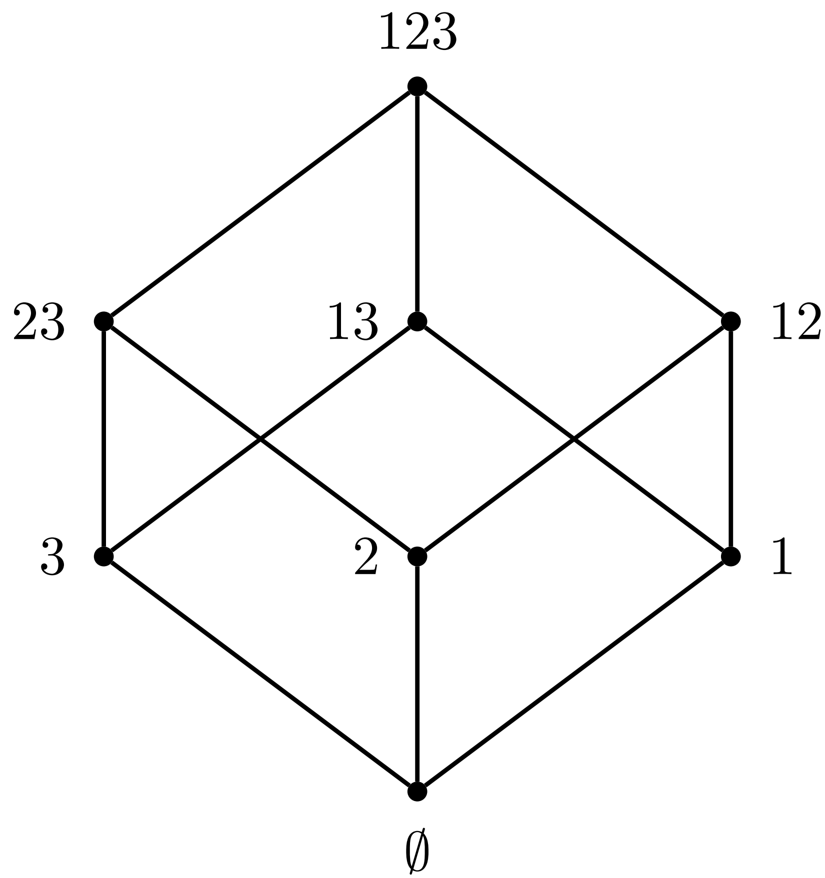
$$I \in \mathcal{F} \;\Leftrightarrow\; \sum_{i \in I} w_i \geq \lambda \tag{1}$$

Axioms guaranteeing the existence of such weights and threshold can possibly be found but they will be hardly interpretable and testable

We choose not to postulate the existence of w_i and λ and to tackle directly the elicitation of $\mathcal{F}$

Highly combinatorial problem

$2^3 = 8$ coalitions



Identifying the set $\mathcal{F}$ of sufficient coalitions $\equiv$ finding a *final* subset of 2^n

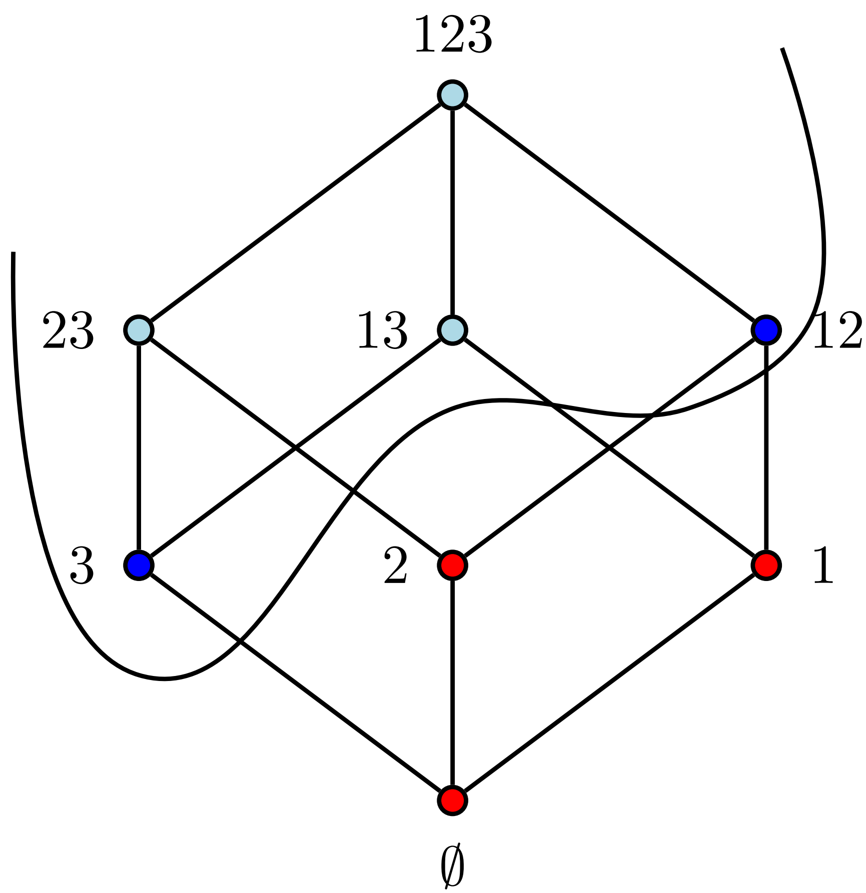
Tantamount to identifying $\mathcal{F}_{min}$ = the set of minimally sufficient coalitions, i.e.

$\mathcal{F}_{min} \subseteq \mathcal{F}$ such that:

- $\forall J \in \mathcal{F},\ \exists I \in \mathcal{F}_{min}$ s.t. $I \subseteq J$ and
- $\forall I, J \in \mathcal{F}_{min},\ I \not\subset J$

The possible $\mathcal{F}_{min}$ are the *antichains* of $2^n, \subseteq$

Example



$\mathcal{F} = \{3, 12, 13, 23, 123\}$

$\mathcal{F}_{min} = \{3, 12\}$

The number of antichains in 2^n forms the integer sequence of *Dedekind* numbers (sequence A000372)

For $n = 2$, there are 6 antichains:

$$\emptyset \mid 1 \mid 2 \mid 12 \mid 1,2 \mid \{\ \} \mid$$

These numbers grow *very fast*:

n	$A000372$
3	20
4	168
5	7581
6	7828354
7	2414682040998

Looking for a needle in a hay stack ...

Finding a particular antichain is a combinatorial search problem like e.g.

- detecting a false coin in a set using a balance
- Master Mind game
- Genomics: reconstruct the right sequencing of pieces of genes $\rightarrow$ problem of finding a particular matching in a graph, or a particular hamiltonian cycle, or more generally, a particular structure

Efficiency of the search:

- False coin: minimal number of weighings
- Master Mind: minimal number of questions
- Genomics: minimal number of tests or experiments, ...

Connections with theory of information

References: V. Grebinski and G. Kucherov [95 to 98], N. Alon et al. [02], N. Alon et V. Asodi [04], etc

Efficient strategy for eliciting $\mathcal{F}$

Type of question considered : $I \in \mathcal{F}$?

Under the guise:

$$(b_I, a_{-I}) \in \mathcal{A} \text{ ?}$$

What is an efficient strategy?

- minimize the number of questions ?
- minimize the cognitive burden on the DM ? $\rightarrow$ raises the question of the *difficulty* of the questions

What about errors in answering the questions? $\rightarrow$ we neglect this issue

Let us concentrate on the number of questions

It is uncertain $\rightarrow$ minimize expected number of questions $\rightarrow$ which distribution ?

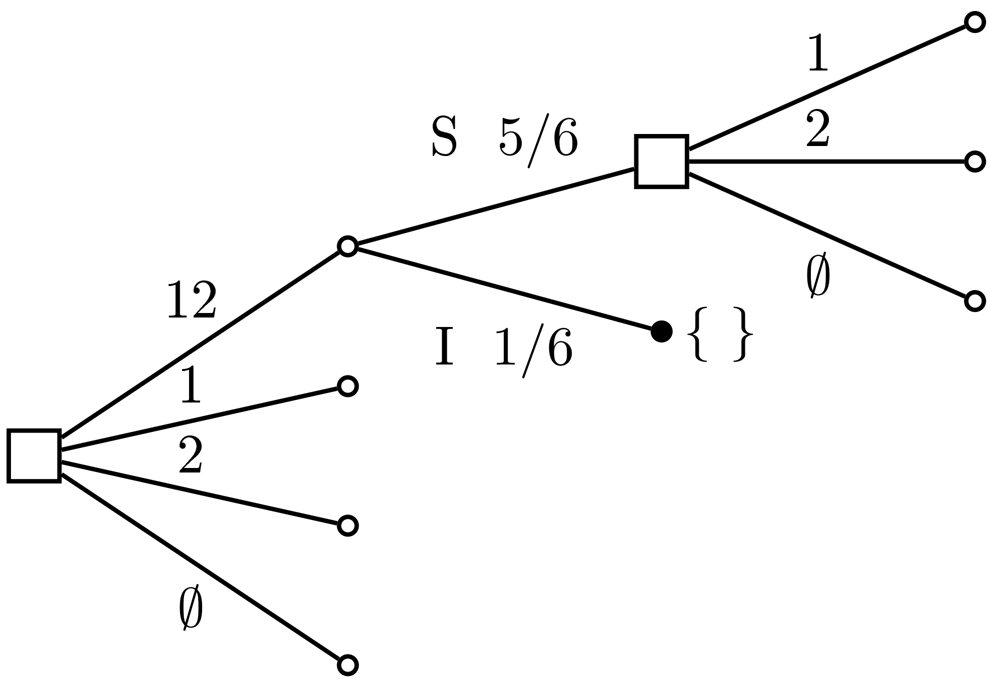
We assume a uniform distribution on the antichains

Exploratory analysis:

- $n = 2$ and $n = 3$
- additional information: $3 \triangleright 2 \triangleright 1$

The case of two criteria

Compute a decision tree for questioning



Conclusion for $n = 2$

On average: $2\frac{2}{3}$ questions

Strategy:

- ask "1 is S ? " (or "2 is S ? ")
- then,
  - if 1 is S, ask "2 is S ?" or "$\emptyset$ is S ? "
  - if 1 is I, ask "2 is S ?" or "12 is S ? "
- etc.

Case $n = 3$ Complicated but feasible by hand; 20 possible $\mathcal{F}_{min}$

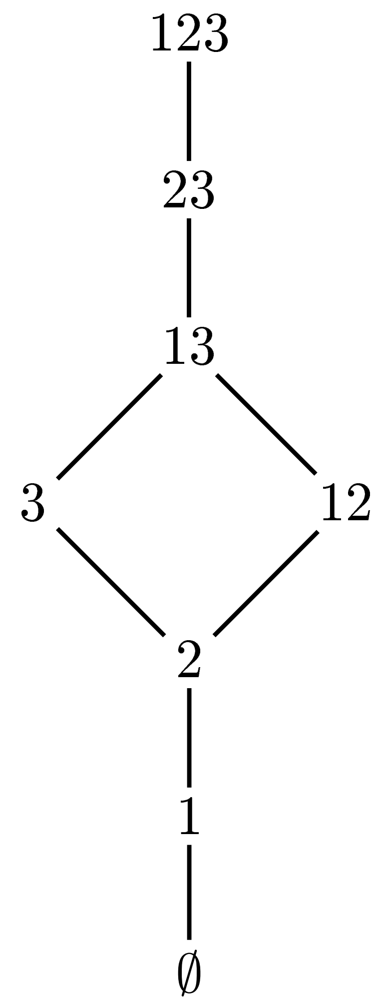
Additional information available

Example: $n = 3$

and $3 \triangleright 2 \triangleright 1$

$\mathcal{F}_{min}$ can be
a singleton
or 3 and 12

10 instead of 20 antichains



Conclusion

Next issues:

- compute the expected number of questions and the questioning strategy for $n = 3, 4$; study the symmetries, the recursiveness, ...
- use (or implement) computer programs to solve the tree for small but larger n
- implement programs to integrate additional information and compute the corresponding decision tree and assess the reduction in the number of questions
- for moderate values of n, combine with other elicitation techniques (find weights compatible with available info (if any) then try to reduce indetermination through questioning)

- how much less expressive are weights w.r.t. sets of sufficient coalitions when n is small ?
- find an upper bound for the number of questions (and an algorithm that stays beyond that bound); we have an algorithm based on depth first search that we suspect (without proof) to cut by two (asymptotically) the number of questions (2^{n-1}). What if additional information is available?

Further issues:

- explore the issue of the cognitive burden on the DM, i.e. take into account the difficulty of the questions in the definition of a questioning strategy
- introduce vetoes
- consider more than two categories