Eliciting the parameters of the Non Compensatory Sorting model

Eda Ersek Uyanık and Marc Pirlot

Université de Mons, Belgium

DA2PL 2026, Brussels
April 16-17, 2026

UMONS

Goal of the talk

- We study the elicitation of the parameters of a method for sorting objects or alternatives into predefined ordered categories ; objects are assessed w.r.t. several criteria
- For instance, we sort objects in two categories : the acceptable and the unacceptable
- The sorting model we deal with is called the Non Compensatory Sorting (NCS) model
- Basically, with two categories, an alternative is acceptable if its performance reaches a certain level on a sufficient coalition of criteria
- The parameters to be elicited are thus the minimal performance on each criterion (limit profile) and all coalitions of criteria that are sufficient

Goal (cont'd)

Elicitation vs Learning

- Elicitation $\approx$ online (active) learning until all parameters are determined
- Several "offline" learning methods have been proposed for learning the parameters of a NCS model on the basis of a given set of assignment examples
- In this work, we consider an ideal situation in which the DM answers queries (assignment examples) as if it were an oracle : he responds without error according to a given NCS model

Motivation

The additive value function model

$x \succsim y$ iff $\sum_i u_i(x_i) \geq \sum_i u_i(y_i)$

- is supported by a theory, i.e., axioms specifying conditions of applicability expressed in terms of the DM's preference $\succsim$;
- this allows to conceive *rigorous* methods for eliciting the model's parameters, e.g., the method of indifference judgments

Motivation

The additive value function model

$x \succsim y$ iff $\sum_i u_i(x_i) \geq \sum_i u_i(y_i)$

- is supported by a theory, i.e., axioms specifying conditions of applicability expressed in terms of the DM's preference $\succsim$;
- this allows to conceive *rigorous* methods for eliciting the model's parameters, e.g., the method of indifference judgments

Outranking methods

- In a family of methods based on principles very different from additive value, Bouyssou and Marchant (2007a,b) have characterized a sorting method called NCS (Non Compensatory Sorting), which is an idealization of Electre Tri
- We investigate rigorous methods for eliciting the parameters of the NCS model in the spirit of what was done for the additive value function model

Summary

- We start by recalling the definition of the NCS model
  For simplicity we concentrate on sorting into two ordered categories and do not consider vetoes
- We present methods for eliciting the limit profile and the set of sufficient coalitions of criteria
  Questions to the DM exclusively consist of assigning well-chosen alternatives to one of the two categories
- We conclude with suggestions of further research directions

The NCS sorting rule (without veto)

The setting

- Set of criteria : $N = \{1, \ldots, n\}$
- Scale of criterion i : X_i endowed with preference order $\geq_i$
- Alternatives : $x = (x_1, \ldots, x_n) \in X = \prod_{i \in N} X_i$
- Categories : acceptable alternatives $\mathcal{A}$;
unacceptable alternatives $\mathcal{U}$
- Assumption : partition $\langle \mathcal{A}, \mathcal{U} \rangle$ compatible with $\geq_i$, for all i

NCS assignment rule

- limit profile : $b \in X$
- family of sufficient coalitions of criteria $\mathcal{F} \subseteq 2^N$

$$x \in \mathcal{A} \quad \text{iff} \quad \{i \in N : x_i \geq_i b_i\} \in \mathcal{F}$$

The set of sufficient coalitions $\mathcal{F}$ in the NCS model

- $\mathcal{F}$ is a set of subsets of criteria with the property : if $I \in \mathcal{F}$ and $I \subseteq J$ then $J \in \mathcal{F}$
- Particular case : MR-Sort model : $\mathcal{F}$ determined by criteria weights w_i and a threshold λ

$$I \in \mathcal{F} \text{ iff } \sum_{i \in I} w_i \geq \lambda$$

- In contrast, NCS allows for modeling criteria interactions
- $\mathcal{F}$ can be represented by a 0-1-*capacity* μ (a *simple game*) :

$$\mu(I) = \begin{cases} 1 & \text{if } I \in \mathcal{F} \\ 0 & \text{if } I \notin \mathcal{F} \end{cases}$$

The set of minimal sufficient coalitions $\mathcal{F}_{\min}$

- Knowing $\mathcal{F}$ is equivalent to knowing $\mathcal{F}_{\min}$
- $\mathcal{F}_{\min}$ is a subset of sufficient coalitions such that removing any criterion from a coalition in $\mathcal{F}_{\min}$ results in an insufficient coalition
- $\mathcal{F}_{\min}$ is an antichain in $2^N, \subseteq$: elements of $\mathcal{F}_{\min}$ are never included in one another

Table of contents

Eliciting the profile b

Assume we know :

- $\mathcal{F}$ thus $\mathcal{F}_{\min}$
- a level M_i that is acceptable and a level m_i that is unacceptable on X_i for all $i \in N$

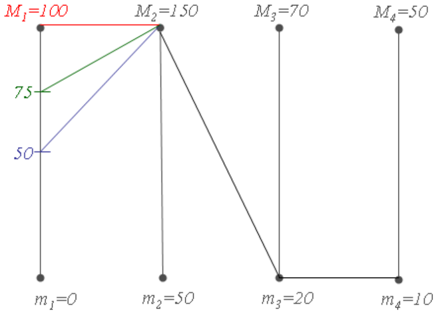
Determining b_i

- Consider a minimal coalition I containing i
- The alternative $(M_I, m_{-I}) \in \mathcal{A}$
- while $(m_i, M_{I\setminus\{i\}}, m_{-I}) \in \mathcal{U}$
- Proceed by **dichotomy** in interval $[m_i, M_i]$

Example

$n = 4$, $\{1,2\} \in \mathcal{F}_{\min}$

$(M_1, M_2, m_3, m_4) \in \mathcal{A}$
$(m_1, M_2, m_3, m_4) \in \mathcal{U}$



Questioning

Q 1 : $(50, M_2, m_3, m_4) \in \mathcal{A}$? Answer 1 : No $\Rightarrow$ $b_i > 50$
Q 2 : $(75, M_2, m_3, m_4) \in \mathcal{A}$? Answer 2 : Yes $\Rightarrow$ $75 \geq b_i > 50$
Q 3 : $(63, M_2, m_3, m_4) \in \mathcal{A}$? ...

Determining b_i

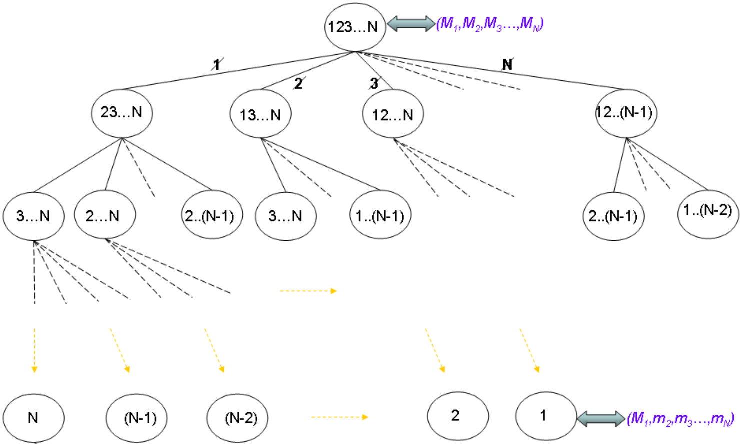
- In case X_i is a finite set, b_i can ideally be precisely determined
- If X_i is infinite or if the DM is uncertain about the answer or if asking more questions is not reasonable, then the process ends up with an interval containing b_i.
- In the example, after 2 questions, we know that $b_i \in]50, 75]$

Eliciting $\mathcal{F}$

Rigorous questions : in terms of alternatives

In practice, for asking whether $I \in \mathcal{F}$,
we ask : Does alternative $(M_I, m_{-I}) \in \mathcal{A}$?

Goal : efficiently explore the set of all coalitions 2^N



Combining Depth First Search and Dichotomy

Initial idea :

- Build the depth first search (DFS) arborescence of the graph of coalitions
- Browse it and Question in the middle of the unknown part of the branches
- Answer : $(M_I, m_{-I}) \in \mathcal{A} \quad \Rightarrow I$ is sufficient (S) ; otherwise, I is insufficient (I)

Example $n = 3$, $\mathcal{F}_{\min} = \{12, 13, 23\}$

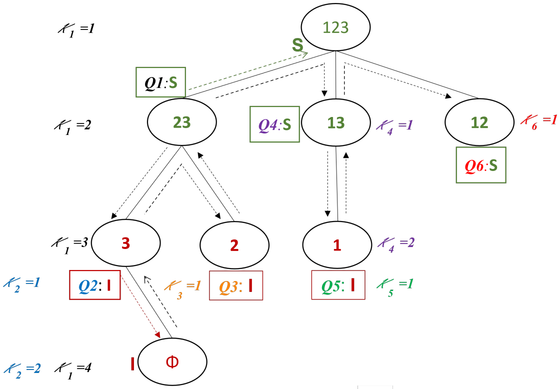


Table of contents

Optimal questioning strategies

- What is a good questioning strategy ?
- Minimizing the number of asked questions
- Two approaches :
  - in the worst case
  - on average

Optimal questioning strategy in the worst case

- The problem was solved by Hansel (1966) in the context of positive Boolean functions
- Hansel's algorithm was studied and used in medical applications by Kovalerchuk et al. (1996), Torvik and Triantaphyllou (2002, 2003)

Optimal questioning strategy in the worst case

- The problem was solved by Hansel (1966) in the context of positive Boolean functions
- Hansel's algorithm was studied and used in medical applications by Kovalerchuk et al. (1996), Torvik and Triantaphyllou (2002, 2003)

Correspondence with positive Boolean functions

- Boolean function : $f : \{0, 1\}^n \to \{0, 1\}$
- positive (= monotone) Boolean function : $f(x) \geq f(y)$ whenever $x \geq y$, where x, y are Boolean vectors
- Correspondence :
  - each Boolean vector represents a subset of N (a coalition) and conversely
  - the set $\mathcal{F}$ of sufficient coalitions corresponds to the set Boolean vectors mapped onto 1 by a positive Boolean function
  - positive Boolean function = 0-1-capacity = simple game

Optimal questioning strategy in the worst case

Principle of Hansel algorithm

- Partition the partially ordered set $\{0,1\}^n, \geq$ into Hansel chains
- Start questioning by the shortest Hansel chains

Optimal questioning strategy in the worst case

Principle of Hansel algorithm

- Partition the partially ordered set $\{0,1\}^n, \geq$ into Hansel chains
- Start questioning by the shortest Hansel chains

Properties of Hansel chains

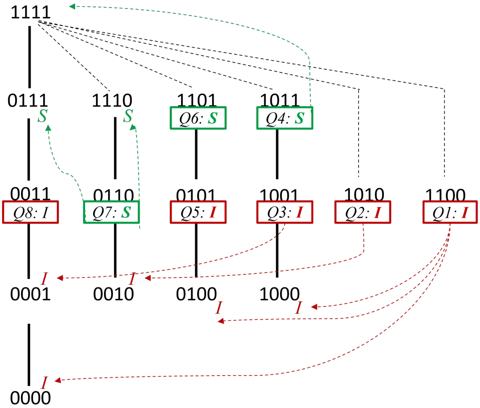
- If we start questioning with the shortest Hansel chains, we ask at most 2 questions on each chain
- In the worst case, Hansel's algorithm asks

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} + \binom{n}{\lfloor \frac{n}{2} \rfloor + 1}$$

questions, which is the lower bound (Korobkov, 1965)

Example $n = 4$

$$\mathcal{F}_{\min} = \{\{2,3\},\{1,2,4\},\{1,3,4\}\} = \{0110, 1101, 1011\}$$



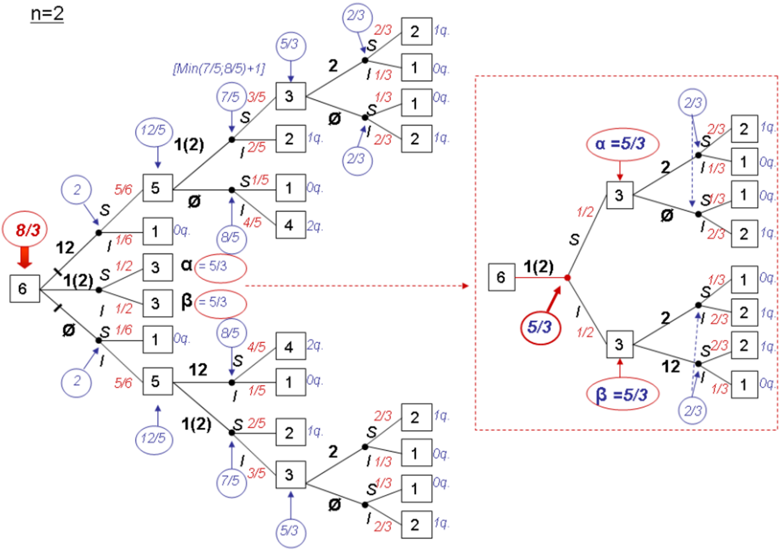
On the construction of Hansel chains

- The chains construction proposed by Hansel is recursive w.r.t. n
- We designed an algorithm for directly obtaining the Hansel chains for n without needing to construct those for $m < n$
- Outline : by browsing through the Hasse diagram of the $\geq$ relation on $\{0,1\}^n$ in an appropriate way, an arborescence is created that contains the Hansel chains ; one can go through this arborescence as done in the Hansel algorithm

Best questioning strategy on average

- Assumption : all $\mathcal{F}_{\min}$ (antichains) are equally likely
- Finding an optimal questioning strategy amounts to solve a decision tree in the uncertain
- Decision tree
  - Decision node, choose a question : is coalition I sufficient ?
  - Chance node : I is sufficient or I is insufficient
- Some tricks to speed up : symmetry, using bounds on the number of questions to prune the tree

Example $n = 2$. Six possible $\mathcal{F}_{\min}$: $\emptyset, 1, 2, \{1, 2\}, 12, \{\}$



Minimal expected number of questions : 8/3

Results

- No general solution is known
- We solved trees by hand up to $n = 4$. For $n = 4$, the optimal average number of questions is $\frac{1286}{168} \approx 7.65$
- A heuristic was proposed and assessed by simulation by Torvik and Triantaphyllou (2002)

Experimental comparison

- We implemented the Depth First Search + Dichotomy algorithm and the Hansel algorithm
- We computed the average number of questions for each of them by applying them to all antichains for $n = 1$ to 6.

$n =$	2	3	4	5	6
D_n : # antichains	6	20	168	7581	7828354
All questions	4	8	16	32	64
Lower bound $\log_2(D_n)$	2.58	4.32	7.39	12.89	22.90
DFS + Dichotomy	2.67	4.65	8.09	15.29	27.94
Hansel	2.67	4.70	7.90	14.50	25.17
Optimal expected	2.67	4.55	7.65	?	?

Comments

- The simulations are exhaustive up to $n = 6$. For $n = 7$, the number of antichains is $D_7 > 2 \cdot 10^{12}$.
- Hansel algorithm shows better performance than DFS+Dichotomy on average
- Torvik and Triantaphyllou (2002) performed exhaustive simulations up to $n = 5$ for Hansel and their heuristic. For values of n between 6 and 11, they make the simulations using an antichains sample[1].
- The heuristic of Torvik and Triantaphyllou performs better than Hansel on average, yet it is not optimal.

1. Though not uniform, their sampling procedure makes a step in the direction of uniform sampling of antichains

Usefulness of DFS+Dichotomy

Is DFS+Dichotomy useless ?

- Requires more questions than Hansel on average, but has an advantage
- DFS+Dichotomy can be applied in case additional information is available
- For instance, information about the relative importance of the criteria
- Hansel only applies to the graph of inclusion of subsets of N ; it may thus loose its advantage

Example

$n = 4$

order of importance

$4 \trianglerighteq 3 \trianglerighteq 2 \trianglerighteq 1$

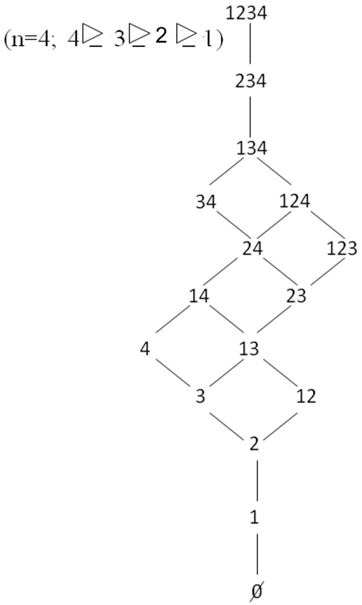
and

$\trianglerighteq$ "propagates "

to coalitions

e.g., [13 is S] $\Rightarrow$ [23 is S]

and [14 is S]



Example

$n = 4$

$4 \trianglerighteq 3 \trianglerighteq 2 \trianglerighteq 1$

and

$\trianglerighteq$ "propagates "
to coalitions

Let $\mathcal{F}_{\min} = \{34, 123, 124\}$

- DFS+Dichotomy (adapted graph) : 5 q.
- DFS+Dichotomy+addinfo : 6 q.
- Hansel+addinfo : 8 q.

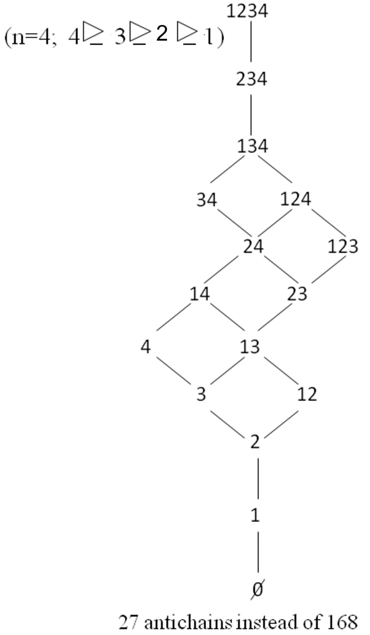


Table of contents

Further work to be done

- Investigate further the computation of the average number of questions for eliciting $\mathcal{F}$
- Find an efficient method for eliciting the limit profile and the sufficient coalitions simultaneously
- Sorting into several categories
  - Torvik and Triantaphyllou (2003) : nested Boolean functions
  - elicit several limit profiles and the corresponding sufficient coalitions
- NCS with veto : elicitation of veto levels
- Usefulness of what has been done for the practical elicitation or learning of NCS and Electre Tri

Thank you for your attention

D. Bouyssou and T. Marchant. An axiomatic approach to noncompensatory sorting methods in MCDM, I : The case of two categories. *European Journal of Operational Research*, 178(1) :217–245, 2007a.

D. Bouyssou and T. Marchant. An axiomatic approach to noncompensatory sorting methods in MCDM, II : More than two categories. *European Journal of Operational Research*, 178(1) :246–276, 2007b.

E. Ersek Uyanık, O. Sobrie, V. Mousseau, and M. Pirlot. Families of sufficient coalitions of criteria involved in ordered classification procedures. Submitted, 2016.

G. Hansel. Sur le nombre de fonctions booléennes de n variables. *Comptes-Rendus de l'Académie des Sciences de Paris (Série A)*, 262 (20) :1088–1090, 1966.

B.K. Korobkov. O monotonnych funkcijach algebry logiki. *Problemy Kibernetiki*, 13 :5–28, 1965.

B. Kovalerchuk, E. Triantaphyllou, A. S. Deshpande, and E. Vityaev. Interactive learning of monotone boolean functions. *Inf. Sci.*, 94(1-4) : 87–118, 1996.

B. Torvik and E. Triantaphyllou. Minimizing the average query complexity of learning monotone Boolean functions. *INFORMS Journal* *on Computing*, 14(2) :142–172, 2002.

B. Torvik and E. Triantaphyllou. Guided inference of nested monotone Boolean functions. *Information Sciences*, 151 :171–200, 2003.

Extension to more than 2 categories

Problem : sorting in p categories (no veto)

- elicit $p-1$ limit profiles $b^1, \ldots, b^{h-1} \ldots b^{p-1}$ with $b_i^h \geq_i b_i^{h-1}$ for all i and $h = 2, \ldots, p-1$
- elicit $\mathcal{F}^h$, for $h = 1, \ldots, p-1$, the sets of sufficient coalitions at level h, with $\mathcal{F}^h \subseteq \mathcal{F}^{h-1}$

Simpler case : $\mathcal{F}^h = \mathcal{F}$ for all h

Extension to more than 2 categories

Problem : sorting in p categories (no veto)

- elicit $p-1$ limit profiles $b^1, \ldots, b^{h-1} \ldots b^{p-1}$ with $b_i^h \geq_i b_i^{h-1}$ for all i and $h = 2, \ldots, p-1$
- elicit $\mathcal{F}^h$, for $h = 1, \ldots, p-1$, the sets of sufficient coalitions at level h, with $\mathcal{F}^h \subseteq \mathcal{F}^{h-1}$

Simpler case : $\mathcal{F}^h = \mathcal{F}$ for all h

Idea : sort iteratively in 2 categories

Inspired byTorvik and Triantaphyllou (2003) : inference of nested Boolean functions

- elicit parameters of bipartition $\langle C^1, \cup_{h=2}^{h=p} C^h \rangle$
- $\Rightarrow$ elicit b^1 and $\mathcal{F}^1$
- Then elicit b^2, the limit profile of the bipartition $\langle C^1 \cup C^2, \cup_{h=3}^{h=p} C^h \rangle$
- and elicit $\mathcal{F}^2$ by determining the coalitions in $\mathcal{F}^1$ that are not sufficient at level 2

NCS 2 categories with veto

- Additional parameter : veto level v_i on criterion i for all or some criteria
- $v_i \leq_i b_i$
- if there is no veto on i, we say by convention that $x_i \geq_i v_i$, for all $x_i \in X_i$

The NCS assignment rule with veto

$$x \in \mathcal{A} \quad \text{iff} \quad \{i \in \mathcal{N} : x_i \geq_i b_i\} \in \mathcal{F} \quad \text{and} \quad x_i \geq v_i \text{ for all } i$$

Is there a veto threshold on i ?

Query : Does $(m_i, M_{-i}) \in \mathcal{A}$?

- Yes $\Rightarrow$ no veto on i and $N \setminus \{i\} \in \mathcal{F}$
- No $\Rightarrow$ two possible interpretations :
  - $N \setminus \{i\} \notin \mathcal{F}$ (which implies that all sufficient coalitions contain i)
    moreover it is uncertain whether there is a veto on i or not
  - $N \setminus i \in \mathcal{F} \Rightarrow$ there is a veto on i

Is there a veto threshold on i ?

Query : Does $(m_i, M_{-i}) \in \mathcal{A}$?

- Yes $\Rightarrow$ no veto on i and $N \setminus \{i\} \in \mathcal{F}$
- No $\Rightarrow$ two possible interpretations :
  - $N \setminus \{i\} \notin \mathcal{F}$ (which implies that all sufficient coalitions contain i)
    moreover it is uncertain whether there is a veto on i or not
  - $N \setminus i \in \mathcal{F} \Rightarrow$ there is a veto on i

Assuming $(m_i, M_{-i}) \notin \mathcal{A}$

Let q_i be the smallest value s.t. $(q_i, M_{-i}) \in \mathcal{A}$

- Difficult to tell whether
  - $q_i = v_i$
  - or $q_i = b_i$
- Yet this is crucial for the determination of the sufficient coalitions : we need a value not in the veto zone but is $< b_i$

In practice

- Usually, fully rigorous elicitation procedures cannot be used in practice
- Can they help in designing procedures that depart as little as possible from rigorous ones ?

In practice

- Usually, fully rigorous elicitation procedures cannot be used in practice
- Can they help in designing procedures that depart as little as possible from rigorous ones ?

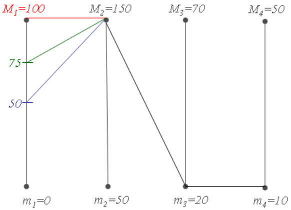
Some perspectives

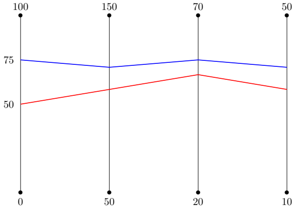
- Finding the sufficient coalitions (or an approximation of $\mathcal{F}$) when n is large
- Managing the imprecise determination of the limit profile

Eliciting $\mathcal{F}$ for $n \geq 6$

- For $n = 6$, the number of questions on average is about 25 ; in the worst case, it is 35
- For $n \geq 6$,
  - seek for more information, e.g. relative importance of criteria
  - organize the criteria into a hierarchy (Kovalerchuk et al. (1996))
  - In case the elicitation process of $\mathcal{F}$ get stuck, one can use linear programming to find weights or a k-additive capacity that represent the available information (Ersek Uyanık et al., 2016)

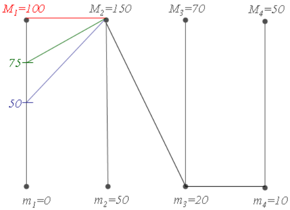
Imprecise determination of the limit profile

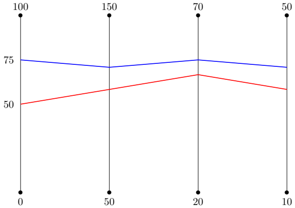




- The DM may be unable to precisely specify the profile values ; one may be left with an interval of possible profile values on each criterion
- Robust assignment : alternative *necessarily* (resp. *possibly)* assigned to category $\mathcal{A}$ if it outranks the upper (resp. lower) limit profile

Imprecise determination of the limit profile

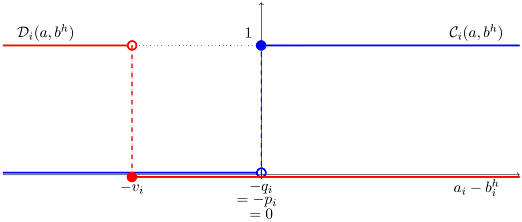




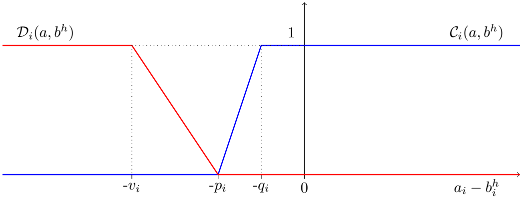
- The DM may be unable to precisely specify the profile values ; one may be left with an interval of possible profile values on each criterion
- Robust assignment : alternative *necessarily* (resp. *possibly)* assigned to category $\mathcal{A}$ if it outranks the upper (resp. lower) limit profile
- Same procedure could be used to elicit the "locally compensatory" interval in Electre Tri

Differences between NCS and Electre Tri

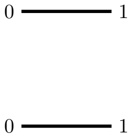
NCS : outranking relation of Electre I



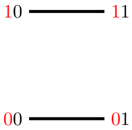
Electre Tri : outranking relation of Electre III



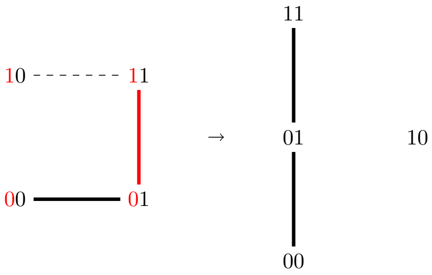
Hansel chains $n = 2$



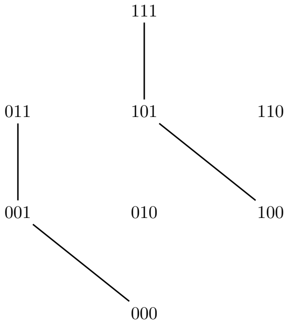
Hansel chains $n = 2$



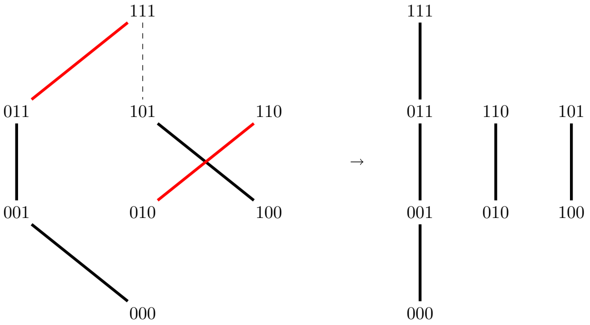
Hansel chains $n = 2$



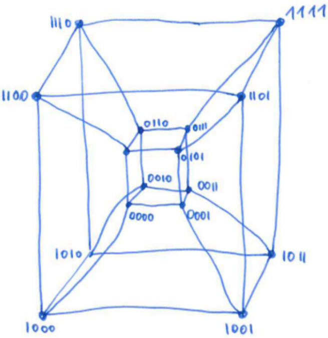
Hansel chains $n = 3$



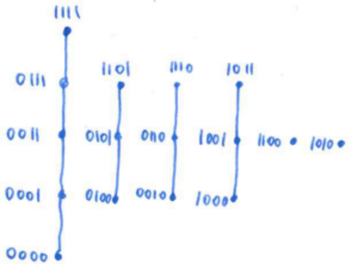
Hansel chains $n = 3$



Hansel chains $n = 4$



Hansel chains $n = 4$



Number of Hansel chains

The Hansel chains provide a decomposition of the n-dimensional cube in disjoint chains :

- 1 chain of length $n + 1$
- $\binom{n}{1} = n$ chains of length $\geq n - 1$
- $\binom{n}{2} = n(n-1)/2$ chains of length $\geq n - 3$
- ...
- $\binom{n}{\lfloor n/2 \rfloor}$ chains of length $\geq n - 2\lfloor n/2 \rfloor + 1 = 1$ or 2

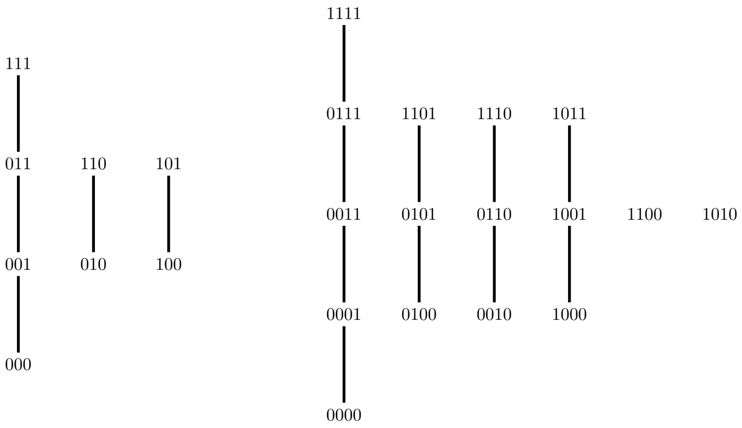
Number of Hansel chains

The Hansel chains provide a decomposition of the n-dimensional cube in disjoint chains :

- 1 chain of length $n+1$
- $\binom{n}{1} = n$ chains of length $\geq n-1$
- $\binom{n}{2} = n(n-1)/2$ chains of length $\geq n-3$
- $\dots$
- $\binom{n}{\lfloor n/2 \rfloor}$ chains of length $\geq n - 2\lfloor n/2 \rfloor + 1 = 1$ or 2

i.e., $\binom{n}{p}$ chains of length $\geq n - 2p + 1$, for $p = 1, \dots, \lfloor n/2 \rfloor$.

Examples

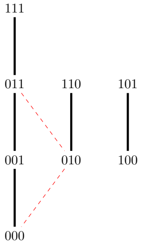


$n = 3$

$n = 4$

The rectangle property

Any three consecutive elements of a chain of length $n - 2p + 1$ form a “rectangle” with an element of a chain of length $n - 2p - 1$



$n = 3$

If we know the class of 010 (I or S), this implies the class of either 000 or 010

Hansel algorithm

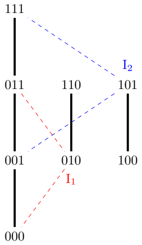
If we know the class (I or S) of all elements in the chains of length $n - 2p - 1$

Hansel algorithm

If we know the class (I or S) of all elements in the chains of length $n - 2p - 1$
only 2 questions are needed for determining the class of each element of a chain of length $n - 2p + 1$

Hansel algorithm

If we know the class (I or S) of all elements in the chains of length $n-2p-1$
only 2 questions are needed for determining the class of each element of a chain of length $n-2p+1$



Example : if 001 and 000 are Insufficient
Only 2 questions : 111 and 011 ?

Hansel algorithm

Algorithm

- Start questioning with the shorter chains
- Propagate what is known for chains of length $n-2p-1$ to chains of length $n-2p+1$
- Ask the remaining (at most) two questions for each chain of length $n-2p+1$

Torvik and Triantaphyllou (2002)'s experiments