Algorithmic Decision Theory

for solving complex decision problems

Raymond Bisdorff

Computer Science and Communications / ILIAS
Faculty of Sciences, Technology & Communication
University of Luxembourg
http://leopold-loewenheim.uni.lu/bisdorff/

Luxembourg, May 3, 2017

Acknowledgments

This presentation contains ideas that are not only the author's.
They have been borrowed from friends and colleagues :

Denis Bouyssou, Luis Dias,
Claude Lamboray, Patrick Meyer,
Vincent Mousseau, Alex Olteanu,
Marc Pirlot, Thomas Veneziano,
and especially,
Alexis Tsoukiàs.

A. Tsoukiàs

Their help is gratefully acknowledged.

Content

1. Introduction to Algorithmic Decision Theory
   - Historical notes
   - Activities and online resources
   - Types of decision problems
2. The Decision Aiding Process
   - Formulating the decision problem
   - Constructing preferential statements
   - The bipolar-valued outranking relation
3. Constructing decision recommendations
   - Selecting
   - Ranking
   - k-Rating / Sorting
   - Clustering

Historical notes : The COST Action IC0602

- From 2007 to 2011 the *Algorithmic Decision Theory* COST Action IC0602, coordinated by Alexis Tsoukiàs, gathered researchers coming from different fields such as Decision Theory, Discrete Mathematics, Theoretical Computer Science and Artificial Intelligence in order to improve decision support in the presence of massive data bases, combinatorial structures, partial and/or uncertain information and distributed, possibly interoperating decision makers.
- Working Groups :
  - Uncertainty and Robustness in Planning and Decision Making
  - Decision Theoretic Artificial Intelligence
  - Preferences in Reasoning and Decision
  - Knowledge extraction and Learning

Historical notes : The CNRS GDRI ALGODEC

- In 2011, the French CNRS, in cooperation with the Belgian FNRS and the FNR, installed a *Groupement de Recherche International* GDRI ALGODEC in order to continue the research on *Algorithmic Decision Theory* by federating a number of international research institutions strongly interested in this research aera.
- The aim is networking the many initiatives undertaken within this domain, organising seminars, workshops and conferences, promoting exchanges of people (mainly early stage researchers), building up an international community in this exciting research area.
- http://www.gdri-algodec.org/

ALGODEC Members

The GDRI ALGODEC was extended 2015 until 2019 and at present involves the following institutions :

DIMACS - Rutgers University (USA)

LAMSADE - Université Paris-Dauphine (FR)

LIP6 - Université Pierre et Marie Curie (Paris, FR)

CRIL - Université d'Artois (Lens, FR)

HEUDIASYC - Université Technologique de Compiègne (FR)

LGI - CentraleSupélec (Paris, FR)

MATHRO - Université de Mons (BE)

SMG - Université Libre de Bruxelles (BE)

ILIAS - University of Luxembourg (LU)

CIG - University Paderborn (DE)

IDSE - Free University Bozen-Bolzana (IT)

GDRI ALGODEC activities

- The International Conferences on *Algorithmic Decision Theory* : ADT'2009 (IT), ADT'2011 (US), ADT'2013 (BE), ADT'2015 (US), ADT'**2017** (LU)
- The workshops DA2PL on *Multiple Criteria Decision Aid and Preference Learning* : 2012 (FR), 2014 (BE) and 2016 (DE)
- The GRAPHS&DECISIONS conference 2014 (LU)
- EURO working groups on *Multiple Criteria Decision Aid* and on *Preference Handling*
- The DIMACS Special Focus on *Algorithmic Decision Theory*
- The International Workshops on *Computational Social Choice*
- *Smart Cities and Policy Ananlytics* Workshops
- The DECISION DECK project

GDRI ALGODEC Online Resources

Tutorials and course materials on http://www.algodec.org.

44 contributions on Algorithmic Decision Theory contain videos and presentation materials originating from the tutorials and courses who took place at the meetings and doctoral schools organised by the COST Action IC0602 Algorithmic Decision Theory.

Types of Decision Problems : Notation

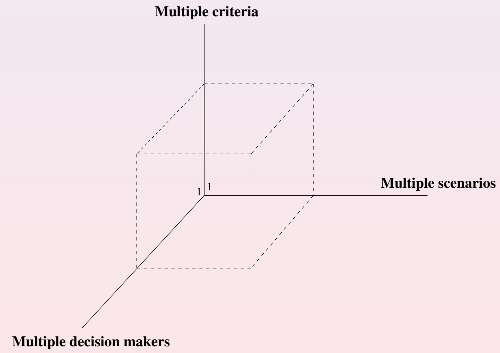
A decision problem will be a tuple $\mathcal{P} = (D, A, O, F, \Omega)$ where

1. D is a group of $d = 1, \dots$ decision makers :
2. A is a set of $n = 2, \dots$ decision alternatives ;
3. O is a set of $o = 1, \dots$ decision objectives ;
4. F is a set of $m = 1, \dots$ attributes or performance criteria (to be maximised or minimised) with respect to decision objective $obj \in O$;
5. Ω is a set of $\omega = 1, \dots, p$ potential states of the world or context scenarios.

Types of Decision Problems – continue

We may distinguish different types of decision problems along three directions :

- Single or multiple objectives/criteria,
- Single or multiple decision makers,
- Single or multiple context scenarios.

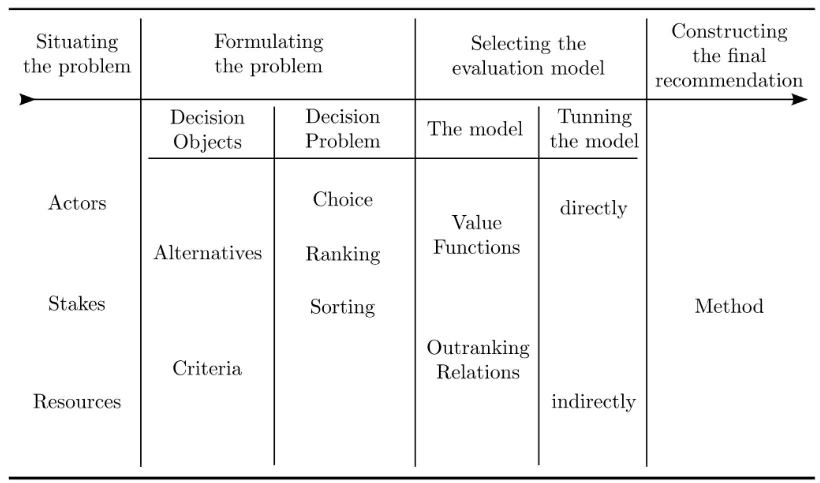


1. Introduction to Algorithmic Decision Theory
Historical notes
Activities and online resources
Types of decision problems

2. The Decision Aiding Process
Formulating the decision problem
Constructing preferential statements
The bipolar-valued outranking relation

3. Constructing decision recommendations
Selecting
Ranking
k-Rating / Sorting
Clustering

Decision aiding process



[Tsoukias:2007]

Formulating the Decision Alternatives

- Small set of individual decision alternatives ;
- Large set of alternatives consisting in the combination of given features ;
- Infinite set of decision alternatives ;
- Portfolios of potential alternatives ;
- Stream of potential decision alternatives ;
- Critical decision alternatives (emergency or disaster recovering).

Formulating decision objectives and criteria

- Identifying the strategic objectives of the decision making problem,
- Identifying all objective consequences of the potential decision actions, measured on :
  - Discrete ordinal scales ?
  - Numerical, discrete or continuous scales ?
  - Interval or ratio scales ?
- Each consequence, measured on a performance criterion, is associated with a strategic objective
  - to be minimized (Costs, environmental impact, energy consumption, etc) ;
  - to be maximised (Benefits, energy savings, security and reliability, etc).
- Verifying the coherence –universal, minimal and separable– of the family of criteria.

Identifying the decision problematique

From an algorithmic point of view, we may distinguish the following decision problematiques :

- Choice : selecting the k best (or worst) choices, $k = 1, \ldots$;
- Weakly ordering : ordering with ties $k = 1, \ldots, n$ choices from a worst to a best equivalence class ;
- Ranking : linearly ordering $k = 1, \ldots, n$ choices from the best to the worst ;
- Sorting/Rating : Supervised clustering into $k = 2, \ldots$ predefined, and usually linearly ordered, sort categories ;
- Relational Clustering : unsupervised grouping into an unknown number $k = 2, \ldots$ of (partially) related clusters.

Example (Adversary or consensual social choice ?)

- Let $\{a, b, c, \ldots, , y, z\}$ be the set of 26 candidates for a 100 voters election. Suppose that :
  - 51 voters have preferences $abc\ldots yz$, and
  - 49 voters have preferences $zbc\ldots ya$.
- 51 voters will vote for a and 49 for z.

Comment

- *In all uninominal election systems, candidate ‘a’ will be elected.*
- *Is ‘a’ really a convincing best candidate ?*
- *No : Nearly half of the voters see candidate ‘a’ as their worst choice !*
- *Whereas candidate ‘b’ could be an unanimous second best candidate !*
- *Simple majority allows dictatorship of majority and does not favor consensual solutions.*

Constructing preferential statements : Notation

- Let X be a finite set of p decision alternatives.
- Let F be a finite set of n criteria (voters) supporting an increasing real performance scale from 0 to M_j ($j = 1, \dots n$).
- Let $0 \leqslant ind_j < pr_j < v_j \leqslant M_j + \epsilon$ represent resp. the indifference, the preference, and the considerable large performance difference discrimination threshold observed on criterion j.
- Let w_j be the significance of criterion j.
- Let W be the sum of all criterion significances.
- Let x and y be two alternatives in X.
- Let x_j be the performance of x observed on criterion j

Performing marginally "*at least as good as*"

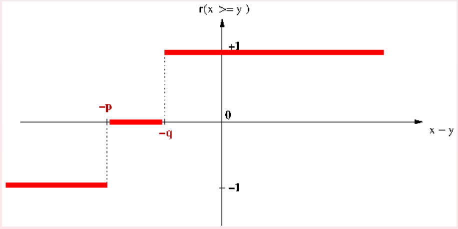
Each criterion j is characterizing a double threshold order $\succsim_j$ on A in the following way :

$$r(x \succsim_j y) = \begin{cases} +1 & \text{if} \quad x_j - y_j \geqslant -ind_j \\ -1 & \text{if} \quad x_j - y_j \leqslant -pr_j \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

+1 signifies x is *performing at least as good as* y on criterion j,

−1 signifies that x is *not performing at least as good as* y on criterion j.

0 signifies that it is *unclear* whether, on criterion j, x is performing at least as good as y.



Performing globally "*at least as good as*"

Each criterion j contributes the significance w_j of his "*at least as good as*" characterisation $r(\succsim_j)$ to the characterisation of a global "*at least as good as*" relation $r(\succsim)$ in the following way :

$$r(x \succsim y) \;=\; \sum_{j \in F} \left[\frac{w_j}{W} \cdot r(x \succsim_j y) \right] \tag{2}$$

$1.0 \geqslant r(x \succsim y) > 0.0$ signifies x is *globally performing at least as good as* y,

$-1.0 \leqslant r(x \succsim y) < 0.0$ signifies that x is *not globally performing at least as good as* y,

$r(x \succsim y) = 0.0$ signifies that it is *unclear* whether x is globally performing at least as good as y.

Performing marginally and globally "*less than*"

Each criterion j is characterising a double threshold order $\prec_j$ (*less than*) on A in the following way :

$$r(x \prec_j y) \quad = \quad \begin{cases} +1 & \text{if} \quad x_j + pr_j \leqslant y_i \\ -1 & \text{if} \quad x_j + ind_j \geqslant y_i \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

And, the *global less than* relation ($\prec$) is defined as follows :

$$r(x \prec y) = \sum_{j \in F} \left[\frac{w_j}{W} \cdot r(x \prec_j y) \right] \tag{4}$$

Property (Coduality principle)

The global "less than" relation $\prec$ *is the dual* ($\not\succsim$) *of the global "at least as good as" relation* $\succsim$.

Considerably *better* or *worse* performing situations

We define a single threshold order, denoted $\ll_j$ which represents *considerably less performing* situations as follows :

$$r(x \ll_j y) = \begin{cases} +1 & \text{if} \quad x_j + v_j \leqslant y_j \\ -1 & \text{if} \quad x_j - v_j \geqslant y_j \\ 0 & \text{otherwise.} \end{cases} \qquad (5)$$

And a codual *considerably better performing* situation $\gg_j$ characterised as :

$$r(x \gg_j y) = \begin{cases} +1 & \text{if} \quad x_j - v_j \geqslant y_i \\ -1 & \text{if} \quad x_j + v_j \leqslant y_i \\ 0 & \text{otherwise.} \end{cases} \qquad (6)$$

Content of the lecture | Introduction | Decision aiding | Recommendations | Bibliography

Global *considerably better* or *considerably worse* *performing* situations

A global *veto*, or *counter-veto* situation is defined as follows :

$$r(x \ll y) = \bigcirc\!\!\!\!\vee_{j \in F}\, r(x \ll_j y) \qquad (7)$$
$$r(x \gg y) = \bigcirc\!\!\!\!\vee_{j \in F}\, r(x \gg_j y) \qquad (8)$$

where $\bigcirc\!\!\!\!\vee$ represents the epistemic disjunction (Bisdorff 1997) or symmetric maximum (Grabisch et al. 2009) operator :

$$r \bigcirc\!\!\!\!\vee r' = \begin{cases} \max(r, r') & \text{if} \quad r \geqslant 0 \wedge r' \geqslant 0, \\ \min(r, r') & \text{if} \quad r \leqslant 0 \wedge r' \leqslant 0, \\ 0 & \text{otherwise.} \end{cases} \qquad (9)$$

Content of the lecture | Introduction | Decision aiding | Recommendations | Bibliography

Characterising *veto* and *counter-veto* situations

1. $r(x \ll y) = 1$ iff there exists a criterion i such that $r(x \ll_i y) = 1$ and there does not exist otherwise any criteria j such that $r(x \gg_j y) = 1$.
2. Conversely, $r(x \gg y) = 1$ iff there exists a criterion i such that $r(x \gg_i y) = 1$ and there does not exist otherwise any criteria j such that $r(x \ll_j y) = 1$.
3. $r(x \gg y) = 0$ if either we observe no considerable performance differences or we observe at the same time, both a considerable positive and a considerable negative performance difference.

Property (Codualty principle)

$r(\not\ll)^{-1}$ *is identical to* $r(\gg)$.

Content of the lecture | Introduction | Decision aiding | Recommendations | Bibliography

The bipolar outranking relation $\succsim$

From an epistemic point of view, we say that :

1. alternative x *outranks* alternative y, denoted $(x \succsim y)$, if
   - 1.1 a significant majority of criteria validates a global outranking situation between x and y, and
   - 1.2 no considerable counter-performance is observed on a discordant criterion,
2. alternative x *does not outrank* alternative y, denoted $(x \not\succsim y)$, if
   - 2.1 a significant majority of criteria invalidates a global outranking situation between x and y, and
   - 2.2 no considerably better performing situation is observed on a concordant criterion.

The outranking concept was originally introduced by **B. Roy**(see Roy, B. (1991). *The outranking approach and the foundations of ELECTRE methods*. Theory and Decision, 31 :49–73).

Polarising the global "*at least as good as*" characteristic

The bipolar-valued characteristic $r(\succsim)$ is defined on X^2 as follows :

$$r(x \succsim y) = r(x \succeq y) \circledvee r(x \not\ll_1 y) \circledvee \ldots \circledvee r(x \not\ll_n y)$$

Properties :

1. $r(x \succsim y) = r(x \succeq y)$ if no considerable positive or negative performance differences between x and y are observed,
2. $r(x \succsim y) = 1.0$ if $r(x \succeq y) \geqslant 0$ and $r(x \gg y) = 1.0$,
3. $r(x \succsim y) = -1.0$ if $r(x \succeq y) \leqslant 0$ and $r(x \ll y) = 1.0$,
4. $\succsim$ is weakly complete : Either $r(x \succsim y) \geqslant 0.0$, or, $r(y \succsim x) \geqslant 0.0$ for $\forall (x \neq y) \in X^2$.

Coherence of the bipolar-valued outranking concept

Property (Coduality Principle)

The dual ($\not\succsim$) of the bipolar outranking relation $\succsim$ is identical to the strict converse outranking $\precnsim$ relation.

Proof : We only have to check the case where $r(x \ll_i y) \neq 0.0$ for all $i \in F$. If $r(x \ll y) \neq 0.0$:

$$\begin{aligned} r(x \not\succsim y) &= -r(x \succsim y) = -\big[r(x \succeq y) \circledvee -r(x \ll y)\big] \\ &= \big[-r(x \succeq y) \circledvee r(x \ll y)\big] \\ &= \big[r(x \not\succeq y) \circledvee -r(x \gg y)\big] \\ &= \big[r(x \prec y) \circledvee r(x \not\gg y)\big] = r(x \precnsim y). \end{aligned}$$

Else, there exist conjointly two criteria i and j such that $r(x \ll_i y) = 1.0$ and $r(x \gg_j y) = 1.0$ such that $r(x \succsim y) = r(x \not\succsim y) = r(x \precnsim y) = 0.0$. □

Bipolar outranking digraphs

Definition

- We denote $\widetilde{G}(X, r(\succsim))$ the bipolar-valued digraph modelled by $r(\succsim)$ on the set X of potential decision alternatives.
- $\widetilde{G}(X, \succsim)$ actually minimizes the sum of the Kendall distances with all marginal -single criterion based- outranking digraphs.
- The average absolute value of the r-valuation is called the epistemic determination of $\widetilde{G}(X, r(\succsim))$.
- We denote $G(X, \succsim)$, the crisp digraph associated with $\widetilde{G}$ where we retain all arcs such that $r(x \succsim y) > 0$, called the associated Condorcet or median cut digraph.
- $G(X, \succsim)$ has usually, except from being trivially reflexive, no other relational properties.

Example (Adversary or consensual social choice ? – continue)

- Let us reconsider the set of 26 candidates for a 100 voters election and suppose again that :
  - 51 voters have preferences *abc...yz*, and
  - 49 voters have preferences *zbc...ya*.
- We may consider both coalitions of voters as two performance criteria -*v*1 and *v*2- with respective significance weights 51 and 49.
- Let us suppose that all the voters are more or less indifferent between candidates of adjacent ranks ; a clear preference appearing first with a rank difference of at leat 3. Furthermore, a considerable rank difference of 25 raises a veto, resp. counter-veto situation.
- We may now compute the corresponding bipolar outranking relation.

1. Introduction to Algorithmic Decision Theory
Historical notes
Activities and online resources
Types of decision problems

2. The Decision Aiding Process
Formulating the decision problem
Constructing preferential statements
The bipolar-valued outranking relation

3. Constructing decision recommendations
Selecting
Ranking
k-Rating / Sorting
Clustering

The Best Choice Problematique

- A choice problem traditionally consists in the search for a single best alternative ;
- Pragmatic Best Choice Recommendation - BCR - principles :

  P_1 : Non retainement for well motivated reasons ;
  P_2 : Recommendation of minimal size ;
  P_3 : Stable (irreducible) recommendation ;
  P_4 : Effectively best choice ;
  P_5 : Recommendation maximally supported by the given preferential information.
- The decision aiding process progressively uncovers the best single choice via more and more refined choice recommendations ;
- The process stops when the decision maker is ready to make her final decision.

References : Roy (1991), Bisdorff, Meyer & Roubens (2008).

Useful choice qualifications

Let Y be a non-empty subset of X, called a choice.

- Y is said to be outranking (resp. outranked) iff $x \notin X \Rightarrow \exists y \in Y : r(y \succsim x) > 0 \quad (resp. r(x \succsim y))$.
- Y is said to be independent iff for all $x \neq y$ in Y we have $r(x \succsim y) \leqslant 0$).
- Y is called an outranking kernel (resp. outranked kernel) iff it is an outranking (resp. outranked) and indendent choice.
- Y is called an outranking hyperkernel (resp. outranked hyperkernel) iff it is an outranking (resp. outranked) choice which consists of independent chordless circuits of odd length $\geqslant 1$.

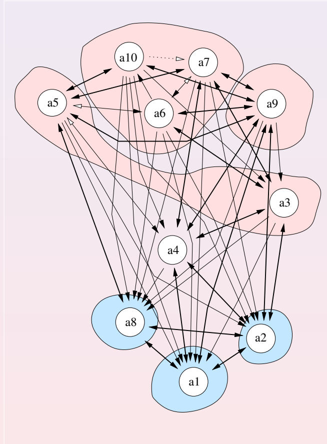
Translating BCR principles into choice qualifications

P_1 : Non-retainment for well motivated reasons.
A BCR is an outranking choice.

P_{2+3} : Minimal size & stable.
A BCR is a hyperkernel.

P_4 : Effectivity.
A BCR is a stricly more outranking than outranked choice.

P_5 : Maximal credibility.
A BCR has maximal determinateness.



Property (BCR Decisiveness, Bisdorff et al. 2008)

Any bipolar strict outranking digraph contains at least one outranking and one outranked hyperkernel.

Adversary or consensual social choice? – continue

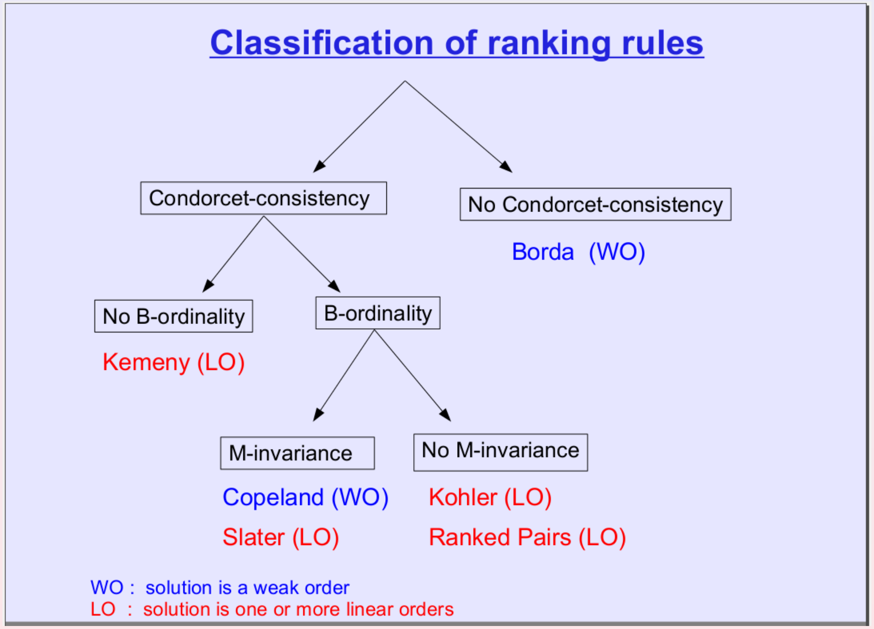
Python3 console session with Digraph3 software resources :

```
>>> from outrankingDigraphs import *
>>> t = PerformanceTableau('electionExample')
>>> g = BipolarOutrankingDigraph(t)
>>> g.showRubisBestChoiceRecommendation()
Rubis best choice recommendation(s) (BCR)
(in decreasing order of determinateness)
Credibility domain: [-100.00,100.00]
=== >> potential best choice(s)
*choice              : [a, b, c]
  +-irredundancy     : 2.00,    independence       : 2.00
  dominance          : 2.00,    absorbency         : -100.00
  covering (%)       : 94.20,   determinateness (%) : 93.37
   - most credible choice(s) = { b: 100.00, c: 49.00, a: 2.00, }
=== >> potential worst choice(s)
* choice             : [x, y, z]
  +-irredundancy     : 2.00,    independence       : 2.00
  dominance          : -100.00, absorbency         : 2.00
  covering (%)       : 0.00,    determinateness (%) : 93.37
  - most credible choice(s) = { y: 100.00, x: 49.00, z: 2.00, }
Execution time: 0.049 seconds
```

Using the Rubis BCR Solver on the leopold-loewenheim.uni.lu server

The Ranking Problem

- A ranking problem traditionally consists in the search for a linear ordering of the set of alternatives;
- A particular ranking is computed with the help of a ranking rule which aggregates preferences over all decision makers and/or criteria into a global (weak) order based, either on (rank) scoring (Borda), or, on (pairwise) voting procedures (Kemeny, Slater, Copeland, Kohler, Ranked Pairs);
- Characteristic properties of ranking rules :
  1. A ranking rule is called Condorcet-consistent when the following holds : If the Condorcet majority relation is a linear order, then this linear order is the unique solution of the ranking rule;
  2. A ranking rule is called B-ordinal if its result only depends on the order of the determination of the r-characteristics B;
  3. A ranking rule is called M-invariant if its result only depends on the median-cut Condorcet relation M.



Ranking of the election candidates

References : Cl. Lamboray (2007,2009), Dias L.C. & Cl. Lamboray (2010)

The k-Rating Problem

- A k-rating problem consists in a supervised partitioning of the set of alternatives into $k = 2, \ldots$ ordred categories, for instance quantile equivalence classes.
- Usually, a rating procedure is designed to deal with a normed evaluation model, whereas choice and ranking algorithms essentially rely on relative evaluation models.
- A crucial problem, hence, lies in the definition of the given categories, i.e., of the *evaluation norms* that define each sort category.
- Two kinds of such norms may be provided :
  - Delimiting norms directly provided by or indirectly learned from observed decision practice;
  - Order statistical computation of quantiles equivalence class delimitations from the corresponding performance tableau.

Performance quantiles

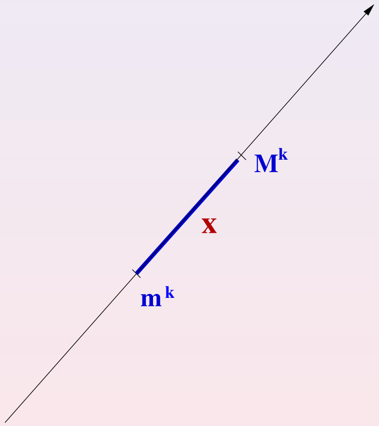
- Let X be the set of n potential decision alternatives evaluated on a single real performance criteria.
- We denote $x, y, \ldots$ the performances observed of the potential decision actions in X.
- We denote quantile $q(p)$ the performance such that $p\%$ of the observed n performances in X are less or equal to $q(p)$.
- The quantile $q(p)$ may be estimated by linear interpolation from the cumulative distribution of the performances in X.

Performance quantile classes

- We consider a series : $p_k = k/q$ for $k = 0, \ldots q$ of $q + 1$ equally spaced quantiles limits like
  - quartiles limits : $0, .25, .5, .75, 1$,
  - quintiles limits : $0, .2, .4, .6, .8, 1$,
  - deciles limits : $0, .1, .2, \ldots, .9, 1$, etc
- The upper-closed q^k class corresponds to the interval $]q(p_{k-1}); q(p_k)]$, for $k = 2, \ldots, q$, where $q(p_q) = \max_X x$ and the first class gathers all data below p_1 : $]-\infty; q(p_1)]$.
- The lower-closed q_k class corresponds to the interval $[q(p_{k-1}); q(p_k)[$,for $k = 1, \ldots, q-1$, where $q(p_0) = \min_X x$ and the last class gathers all data above $q(p_{q-1})$: $[q(p_{q-1}), +\infty[$.
- We call q-tiles a complete series of $k = 1, \ldots, q$ upper-closed q^k, resp. lower-closed q_k, quantile classes.

q-tiles sorting on a single criteria

If x is a measured performance, we may distinguish three sorting situations :



1. $x \leqslant q(p_{k-1})$ and $x < q(p_k)$
   The performance x is lower than the q^k class;
2. $x \geqslant q(p_{k-1})$ and $x < q(p_k)$
   The performance x belongs to the q^k class;
3. $(x > q(p_{k-1})$ and$)$
   $x \geqslant q(p_k)$
   The performance x is higher than the p^k class.

If the relation $<$ is the dual of $\geqslant$, it will be sufficient to check that both, $x \geqslant q(p_{k-1})$, as well as $x \ngeqslant q(p_k)$, are verified for x to be a member of the k-th q-tiles class.

Multiple criteria q-tiles sorting with bipolar outrankings

Property

The bipolar characteristic of x belonging to upper-closed q-tiles class q^k, resp. lower-closed class q_k, may hence, in an outranking approach, be assessed as follows :

$$r(x \in q^k) = \min\left[\, -r\big(\mathbf{q}(p_{k-1}) \succsim x\big),\ r\big(\mathbf{q}(p_k) \succsim x\big) \right]$$
$$r(x \in q_k) = \min\left[\, r\big(x \succsim \mathbf{q}(p_{k-1})\big),\ -r\big(x \succsim \mathbf{q}(p_k)\big) \right]$$

where, for $k = 1, \ldots q$, $\mathbf{q}(p_{k-1})$ and $\mathbf{q}(p_k)$ represent the multiple criteria performance delimitations of quantile q^k, resp. q_k.

Comment

The bipolar outranking relation $\succsim$, verifying actually the coduality principle, $-r\big(\mathbf{q}(p_{k-1}) \succsim x\big) = r\big(\mathbf{q}(p_{k-1}) \not\precsim x\big)$, resp. $-r\big(x \succsim \mathbf{q}(p_k)\big) = r\big(x \not\precsim \mathbf{q}(p_k)\big)$.

The multicriteria (upper-closed) q-tiles sorting algorithm

1. **Input** : a set X of n objects with a performance table on a family of m criteria and a set $\mathcal{Q}$ of $k = 1, .., q$ empty q-tiles equivalence classes.
2. **For each** object $x \in X$ **and each** q-tiles class $q^k \in \mathcal{Q}$:
   2.1 $r(x \in q^k) \quad \leftarrow \quad \min\big(-r(\mathbf{q}(p_{k-1}) \succsim x), r(\mathbf{q}(p_k) \succsim x)\big)$
   2.2 if $r(x \in q^k) \geqslant 0$:
   add x to q-tiles class q^k ;
3. **Output** : $\mathcal{Q}$.

Comment

1. *The complexity of the q-tiles sorting algorithm is $\mathcal{O}(nmq)$; linear in the number of decision actions (n), criteria (m) and quantile classes (q).*
2. *As $\mathcal{Q}$ represents a partition of the criteria measurement scales, i.e. the upper limits of the preceding category correspond to the lower limits of the succeeding ones, there is a potential for run time optimization.*

Example of quintiles sorting result

```
>>> from randomPerfTabs import RandomPerformaceTableau
>>> t = RandomPerformanceTableau(numberOfActons=50,
  ...                                        seed=5)
>>> from sparseOutrankingDigraphs import\
  ...                     PreRankedOutrankingDigraph
>>> pr = PreRankedOutrankingDigraph(t,qintiles)
>>> pr.showSorting()
*--- Sorting results in descending order ---*
]0.8 - 1.0]: [a16, a02, a24, a32]
]0.6 - 0.8]: [a01, a02, a06, a09, a10, a13, a16, a18,
              a22, a25, a27, a28, a31, a32, a36, a37,
              a39, a40, a41, a43, a45, a48]
]0.4 - 0.6]: [a01, a03, a04, a05, a07, a08, a09, a10,
              a11, a12, a13, a14, a15, a17, a18, a20,
              a26, a27, a29, a30, a33, a34, a35, a38,
              a42, a43, a44, a45, a46, a47, a49, a50]
]0.2 - 0.4]: [a04, a11, a12, a17, a19, a21, a23,
              a29, a34, a42, a46, a47, a50]
] < - 0.20]: []
```


Properties of q-tiles sorting result

1. *Coherence* : Each object is always sorted into a non-empty subset of adjacent q-tiles classes.
2. *Uniqueness* : If $r(x \in q^k) \neq 0$ for $k = 1, ..., q$, then performance x is sorted into exactly one single q-tiled class.
3. *Separability* : Computing the sorting result for performance x is independent from the computing of the other performances' sorting results.

Comment

The separability property gives us access to efficient parallel processing of class membership characteristics $r(x \in q^k)$ for all $x \in X$ and q^k in $\mathcal{Q}$.

Ordering the q-tiles sorting result

The q-tiles sorting result leaves us with more or less overlapping ordered quantile equivalence classes. For constructing now a linearly ordered q-tiles partition of X, we may apply three strategies :

1. Optimistic : In decreasing lexicographic order of the upper and lower quantile class limits ;
2. Pessimistic : In decreasing lexicographic order of the lower and upper quantile class limits ;
3. Average : In decreasing numeric order of the average of the lower and upper quantile limits.

Example of quintiles decomposing result

Applying the average ordering strategy we obtain the following weak ordering of the decision alternatives :

```
>>> pr.showDecomposition()
*--- quantiles decomposition in decreasing order---*
    c1. ]0.8-1.0]: [a24]
    c2. ]0.6-1.0]: [a16,a22, a32]
    c3. ]0.6-0.8]: [a02, a06, a25, a28, a31, a36, a37,
                    a39,  a40, a41, a48]
    c4. ]0.4-0.8]: [a01, a09, a10, a13, a18,
                    a27, a43, a45]
    c5. ]0.4-0.6]: [a03, a05, a07, a08, a14, a15, a20,
                    a26, a30, a33, a35, a38, a44, a49']
    c6. ]0.2-0.6]: [a04,'a11, a12, a17, a29, a34, a42,
                    a46, a47, a50]
    c7. ]0.2-0.4]: [a19, a21, a23]
```


q-tiles ranking algorithm

1. **Input** : the outranking digraph $\tilde{\mathcal{G}}(X, \succsim)$, a partition P_q of k linearly ordered decreasing parts of X obtained by the q-tiles sorting algorithm, and an empty list $\mathcal{R}$.
2. **For each** quantile class $q^k \in P_q$:
   if $\#(q^k) > 1$:
   $R_k \leftarrow$ **locally rank** q^k in $\tilde{\mathcal{G}}_{|q^k}$
   (if ties, render alphabetic order of action keys)
   else : $R_k \leftarrow q^k$
   append R_k to $\mathcal{R}$
3. **Output** : $\mathcal{R}$

Example

Quintiles sorting & ranking of the 26 candidates from the election example.

q-tiles ranking algorithm – Comments

1. The complexity of the q-tiles ranking algorithm is linear in the number k of components resulting from a q-tiles sorting which contain more than one action.
2. Three tractable local ranking rules – *Copeland*'s, *Net-flows*' and *Kohler*'s rule – of complexity $\mathcal{O}\big((\#q^k)^2\big)$ may be used on each restricted outranking digraph $\tilde{\mathcal{G}}_{|q^k}$.
3. In case of local ties (very similar evaluations for instance), the **local ranking** procedure will render these actions in increasing alphabetic ordering of the action keys.
4. Large scale Monte Carlo simuations with random performance tableaux of different sizes and types show that the Copeland local ranking rule gives the best practical results, both in terms of execution times, as well as in terms of the ordinal correlation with the corresponding global outranking situations.

5-tiled versus standard outranking digraph of order 50

Symbol legend

- ⊤ outranking for certain
- \+ more or less outranking
- ' ' indeterminate
- − more or less outranked
- ⊥ outranked for certain

Sparse digraph *bg* :
\# Actions : 50
\# Criteria : 7
Sorted by : 5-Tiling
Ranking rule : Copeland
\# Components : 7
Minimal order : 1
Maximal order : 15
Average order : 7.1
fill rate : 20.980%
correlation : +0.7563

Multithreading the q-tiles sorting & ranking procedures

1. Following from the separability property of the q-**tiles sorting** of each action into each q-tiles class, the q-sorting algorithm may be safely split into as much threads as are multiple processing cores available in parallel.
2. Furthermore, the **ranking** procedure being local to each diagonal component, these procedures may as well be safely processed in parallel threads on each restricted outranking digraph $\mathcal{G}_{|q^k}$.

HPC q-tiling & ranking records

digraph order	standard model #c.	t_g sec.	τ_g	q-tiled model #c.	t_{bg}	τ_{bg}
1 000	118	6"	+0.88	8	1.6"	+0.83
2 000	118	15"	+0.88	8	3.5"	+0.83
2 500	118	27"	+0.88	8	4.4"	+0.83
10 000				118	7"	
15 000				118	12"	
25 000				118	21"	
50 000				118	48"	
100 000	(size	=	10^{10})	118	2'	(fill rate = 0.077%)
1 000 000	(size	=	10^{12})	118	36'	(fill rate = 0.028%)
1 732 051	(size	=	3×10^{12})	118	2h17'	(fill rate = 0.010%)
2 500 000	(size	=	6.25×10^{12})	118	2h55'	(Spring 2017)

Legend :

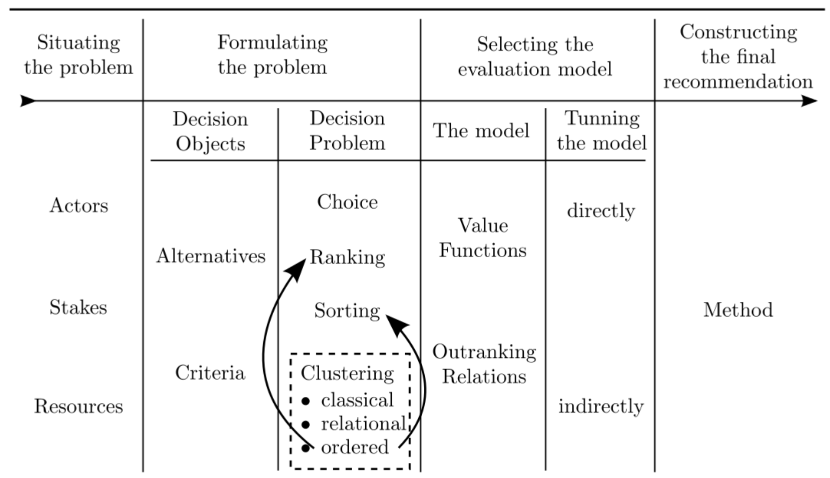
- #c. = number of cores ;
- g : standard outranking digraph, bg : the q-tiled outranking digraph ;
- t_g, resp. t_{bg}, are the corresponding constructor run times ;
- τ_g, resp. τ_{bg} are the ordinal correlation of the Copeland ordering with the given outranking relation.

The clustering problematique

- Clustering is an unsupervised learning method that groups a set of objects into clusters.
- Properties :
  - Unknown number of clusters ;
  - Unknown characteristics of clusters ;
  - Only the relations between objects are used ;
  - no relation to external categories are used.
- Usually used in exploratory analysis and for cognitive artifacts.

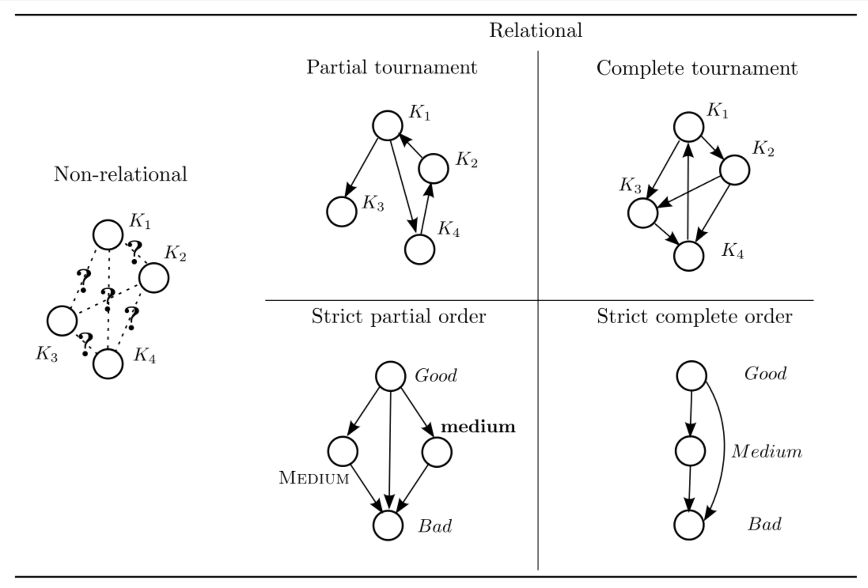
Content of the lecture | Introduction | Decision aiding | Recommendations | Bibliography

Clustering decision aid



[Tsoukias:2007]

Classification of custering approaches



Algorithmic Approach

- define a **fitness function** for each objective:
  - → maximize indifference relations inside clusters;
  - → maximize preference relations between clusters.
- **Exact**:
  1. enumerate all partitions;
  2. select the best w.r.t. the objective;

  → **exponential** number of partitions.
- **Approximative**:
  - Relational Clustering **[de Smet, Eppe: 2009]**;
  - Multicriteria Ordered Clustering **[Nemery, de Smet: 2005]**;
  - CLIP **[Bisdorff, Meyer, Olteanu: 2012]**;
  - ...

Algorithmic Approach – continue

CLIP (CLustering using Indifferences and Preferences)

1. Grouping on indifferences (**internal**);
   - finding an **initial partition**;
   - high concentration of indifference relations inside clusters;
   - low concentration of indifference relations between clusters;
   - graph theoretic inspired method using cluster cores;
2. Refining on preferences (**external**);
   - **searching** for the **optimal result**;
   - strengthen relations between clusters;
   - **meta-heuristic** approach.

Concluding ...

The outranking approach allows to :

- take into account imprecise performance measures via discrimination thresholds ;
- avoid any majority dictatorship by taking into account considerable large negative and positive performance differences ;
- deal with incommensurable and missing data due to a bipolar-valued characteristic encoding of outranking assertions ;
- coherently deal with positive and negative epistemic and logical facts due to the coduality principle ;
- avoid a value theoretic approach and by the way its necessary transitivity axiom of preferences (no need for any kind of dictator ;-) ;
- tackle coherently incomparabilities, cycles, partial and intransitive preferences (no so-called preference paradoxes) ;
- profit from the separability property of the pairwise outranking model for implementing efficient HPC algorithms.

Bibliography I

- R. Bisdorff, L.D. Dias, P. Meyer, V. Mousseau and M. Pirlot, *Evaluation and Decision Models : case studies*. Springer Verlag, Berlin, pages 673, 2015.
- R.I. Brafman, F.S. Roberts and A. Tsoukiàs, *ADT'2011*. Proceedings of the 2nd International Conference on Algorithmic Decision Theory, Springer Verlag, Berlin, 2011.
- F. Rossi and A. Tsoukiàs, *Algorithmic Decision Theory*. Proceedings of the 1st International Conference ADT2009, LNAI 5783, Springer Verlag, Berlin, 460 pages, 2009.
- F.S. Roberts and A. Tsoukiàs, *Special Issue on Voting and Preference Modelling*. Mathematical Social Sciences, vol. 57, 118 pages, 2009.
- F.S. Roberts and A. Tsoukiàs, *Special Issue on Computer Science and Decision Theory*. Annals of Operations Research, vol. 163, 270 pages, 2008.

Bibliography II

- D. Bouyssou, Th. Marchant, M. Pirlot, A. Tsoukiàs and Ph. Vincke, *Evaluation and Decision Models : stepping stones for the analyst*. Springer Verlag, Berlin, 2006.
- P. Perny and A. Tsoukiàs, *Decision analysis and artificial intelligence*. European Journal of Operational Research, vol. 160(3), 300 pages, 2005.
- D. Bouyssou, Th. Marchant, P. Perny, M. Pirlot, A. Tsoukiàs and Ph. Vincke, *Evaluation and Decision Models : a critical perspective*. Kluwer Academic, Dordrecht, 2000.
- R.Bisdorff, *Computing linear rankings from trillions of pairwise outranking situations*. In Proceedings of DA2PL'2016 *From Multiple Criteria Decision Aid to Preference Learning*, R. Busa-Fekete, E. Hüllermeier, V. Mousseau and K. Pfannschmidt (Eds.), University of Paderborn (Germany), Nov. 7-8 2016 : 1-6

Bibliography III

- R. Bisdorff, *On confident outrankings with multiple criteria of uncertain significance*. In Proceedings of DA2PL'2014 *From Multiple Criteria Decision Aid to Preference Learning*, V. Mousseau and M. Pirlot (Eds.), Ecole Centrale Paris, Nov. 20-21 2014 : 119-124
- R. Bisdorff, P. Meyer and Th. Veneziano. *Elicitation of criteria weights maximising the stability of pairwise outranking*. Journal of Multi-Criteria Decision Analysis (Wiley) (2014) 21 : 113-124
- R. Bisdorff, *On polarizing outranking relations with large performance differences*. Journal of Multi-Criteria Decision Analysis (2013) 20 :3-12
- R. Bisdorff, *On measuring and testing the ordinal correlation between bipolar outranking relations*. In Proceedings of DA2PL'2012 *From Multiple Criteria Decision Aid to Preference Learning*, M. Pirlot and V. Mousseau (Eds.), University of Mons, November 15-16 2012, pp. 91-100.

Bibliography IV

- R. Bisdorff, P. Meyer and A. Olteanu, *A Clustering Approach using Weighted Similarity Majority Margins*. Tang et al. (Eds.) : Advanced Data Mining and Applications ADMA 2011, Part I, Springer-Verlag LNAI 7120 (2011) 15–28.
- L.C. Dias and C. Lamboray, *Extensions of the prudence principle to exploit a valued outranking relation*. European Journal of Operational Research Volume 201 Number 3 (2010) 828-837.
- C. Lamboray, *A prudent characterization of the Ranked Pairs Rule*. Social Choice and Welfare 32 (2009) 129-155.
- A. Tsoukiàs, *From Decision Theory to Decision Aiding Methodology*. European Journal of Operational Research, vol. 187 (2008) 138 - 161.

Bibliography V

- R. Bisdorff, P. Meyer and M. Roubens, *RUBIS : a bipolar-valued outranking method for the choice problem*. 4OR, A Quarterly Journal of Operations Research, Springer-Verlag, Volume 6 Number 2 (2008) 143-165.
- C. Lamboray, *A comparison between the prudent order and the ranking obtained with Borda's, Copeland's, Slater's and Kemeny's rules*. Mathematical Social Sciences 54 (2007) 1-16.
- R. Bisdorff, M. Pirlot and M. Roubens, *Choices and kernels from bipolar valued digraphs*. European Journal of Operational Research, 175 (2006) 155-170.
- R. Bisdorff, *Concordant Outranking with multiple criteria of ordinal significance*. 4OR, A Quarterly Journal of Operations Research, Springer-Verlag, Volume 2 Number 4 (2004) 293-308.