MICS 2 : Algorithmic Decision Theory

Raymond Bisdorff

Emeritus Professor of
Applied Mathematics and Computer Science
University of Luxembourg
http://lepold-loewenheim.uni.lu/bisdorff/

Summer Semester 2020

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

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.

AlgoDec Members

The GDRI AlgoDec was extended 2015 until 2019 and involved 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), ADT'2019 (US)
- The workshops DA2PL on *Multiple Criteria Decision Aid and Preference Learning* : 2012 (FR), 2014 (BE), 2016 (DE), 2018 (PL), and 2020 (IT)
- 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

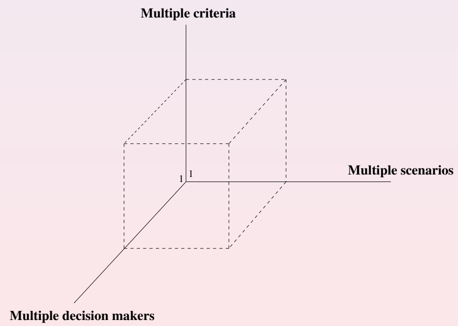
A decision problem will be a tuple $\mathcal{P} = (D, A, O, F, \Omega)$ where

1. D is a group of $d = 1, \ldots$ decision makers :
2. A is a set of $n = 2, \ldots$ decision alternatives ;
3. O is a set of $o = 1, \ldots$ decision objectives ;
4. F is a set of $m = 1, \ldots$ attributes or performance criteria ; each one to be maximised or minimised with respect to a given decision objectives $obj \in O$;
5. Ω is a set of $\omega = 1, \ldots, p$ 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.



Decision aiding process

Timeline : $\longrightarrow$

Situating the problem	Formulating the problem: Decision Objects	Formulating the problem: Decision Result	Selecting the evaluation model: Evaluation model	Selecting the evaluation model: Tuning the parameters	Constructing recommendations
Actors	Objectives	Ranking	Value Functions	directly	Graph kernel extraction
		Choice			
Stakes	Alternatives	Rating	Performance Indicators	indirectly by learning	sorting algorithms
Resources	Performance Criteria	Clustering	Preference modelling		quantiles estimation

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).

Identifying the decision result

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

- Rankings : Sorting the decision alternatives from best to worst ;
- Best Choice : Selecting the k best alternatives, $k = 1, \ldots$;
- Ratings : Supervised sorting of the alternatives into predefined, and usually linearly ordered rating categories ;
- Relational Clusterings : Unsupervised sorting of the alternatives into an unknown number of (partially) related clusters.

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.

Modelling the performance tableau

- Let X be a finite set of p decision alternatives.
- Let F be a finite set of n criteria (voters, experts, ...) supporting an increasing real performance scale from 0 to M_j ($j = 1, \ldots n$).
- 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

Modelling outranking situations

X : Finite set of n alternatives

$x \succsim y$: Alternative x outranks alternative y if

1. there is a (weighted) majority of criteria (voters, experts, ...) supporting that x performs at least as good as y, and
2. no considerable negative performance difference between x and y is observed on a discordant criterion.

$x \not\succsim y$: Alternative x does not outrank alternative y if

1. there is a (weighted) majority of criteria (voters, experts, ...) supporting that x does not perform at least as good as y, and
2. no considerable positive performance difference between x and y is observed on a discordant crterion.

$r(x \succsim y)$ represents a bipolar, i.e. *concordance versus discordance*, valuation in $[-1, 1]$ that characterises the epistemic truth of affirmative assertion $x \succsim y$.

Epistemic truth semantics of the r-valuation

Let $x \succsim y$ and $x' \succsim y'$ be two preferential assertions :

- $r(x \succsim y) = +1$ means that assertion $x \succsim y$ is certainly valid,
- $r(x \succsim y) = -1$ means that assertion $x \succsim y$ is certainly invalid,
- $r(x \succsim y) > 0$ means that assertion $x \succsim y$ is more valid than invalid,
- $r(x \succsim y) < 0$ means that assertion $x \succsim y$ is more invalid than valid,
- $r(x \succsim y) = 0$ means that
  validity of assertion $x \succsim y$ is indeterminate,
- $r(x \succsim y) > r(x' \succsim y')$ means that
  assertion $x \succsim y$ is more valid than assertion $x' \succsim y'$,
- $r(\neg x \succsim y) = -r(x \succsim y)$
  logical (strong) negation by changing sign,
- $r(x \succsim y \vee x' \succsim y') = \max(r(x \succsim y), r(x' \succsim y'))$
  logical disjunction via the *max* operator,
- $r(x \succsim y \wedge x' \succsim y') = \min(r(x \succsim y), r(x' \succsim y'))$
  logical conjunction via the *min* operator.

Coherence of the bipolar outranking concept

Properties :

1. The bipolar outranking relation $\succsim$ is trivially reflexive,
2. The bipolar outranking relation $\succsim$ is weakly complete, ie $r(x \succsim y) < 0$ implies $r(y \succsim x) \geq 0$.
3. The dual ($\not\succsim$) of the bipolar outranking relation $\succsim$ is identical to the strict converse outranking $\precnsim$ relation.

However, other properties, like being *acyclic* or even *transitive* are usually are not fulfilled.

Content of the lecture Introduction Decision aiding Recommendations Bibliography

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 associated Condorcet or median cut digraph, i.e. the crisp digraph associated with $\widetilde{G}$ where we retain all arcs such that $r(x \succsim y) > 0$.
- $G(X, \succsim)$ has usually, except from being trivially reflexive, no other relational properties.

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
   - Selecting the evaluation model
   - Preference modelling via the outranking approach
3. Constructing decision recommendations
   - Constructing Rankings
   - Selecting k-best or -worst choice
   - k-Rating
   - Relational Clustering

Constructing Rankings

Timeline : ⟶

Situating	Formulating the problem		Selecting the evaluation model		Constructing
the problem	Decision Objects	Decision Result	Evaluation Model	Tuning the parameters	recommendations
Actors	Objectives	Ranking	Value Functions	directly	Kernel extraction
		Choice			
Stakes	Alternatives	Rating	Performance Indicators		Sorting algorithms
Resources	Performance Criteria	Clustering	Preference modelling	indirectly by learning	Quantiles estimation

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 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 majority margins B ;
  3. A ranking rule is called M-invariant if its result only depends on the majority relation M.

Reference : Cl. Lamboray (2007,2009,2010)

A classification of ranking rules by Cl. Lamboray

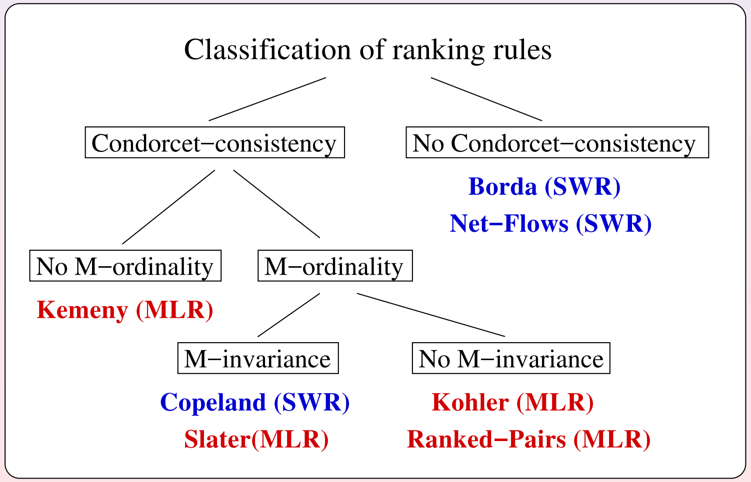


FIGURE – Legend : *SWR* : single weak ranking, *MLR* : multiple linear rankings

Selecting *k*-best or -worst choice

Timeline : $\longrightarrow$

Situating the problem	Formulating the problem		Selecting the evaluation model		Constructing recommendations
	Decision Objects	Decision Result	Evaluation model	Tuning the parameters	
Actors	Objectives	Ranking	Value Functions	directly	Kernel extraction
		Choice			
Stakes	Alternatives	Rating	Performance Indicators		Sorting algorithms
Resources	Performance Criteria	Clustering	Preference modelling	indirectly by learning	Quantiles estimation

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 & Bouyssou (1993), Bisdorff, Roubens & Meyer (2008).

Useful choice qualifications

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

- Y is said to be outranking (resp. outranked) when $x \notin X \Rightarrow \exists y \in Y : r(y \succsim x) > 0 \quad (\textit{resp.}\, r(x \succsim y))$.
- Y is said to be independent (resp. weakly independent) when for all $x \neq y$ in Y we have $r(x \succsim y) < 0)$ (resp. $r(x \succsim y) \leqslant 0)$).
- Y is called an outranking kernel (resp. outranking prekernel) when it is an outranking and indendent (resp. weakly independent) choice.
- Y is called an outranked kernel (resp. outranked prekernel) when it is an outranked and indendent (resp. weakly independent) choice.

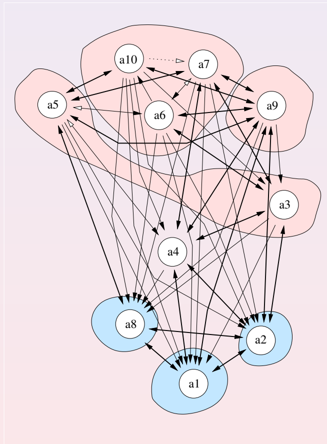
Translating BCR principles into choice qualifications

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

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

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

P_5 : Maximal epistemic support.
A BCR has maximal determinateness.



Property (BCR Decisiveness)

Any bipolar strict outranking digraph without chordless odd circuit contains at least one outranking and one outranked prekernel.

k-Rating

Timeline : $\longrightarrow$

Situating the problem	Formulating the problem: Decision Objects	Formulating the problem: Decision Result	Selecting the evaluation model: Evaluation model	Selecting the evaluation model: Tuning the parameters	Constructing recommendations
Actors	Objectives	Ranking	Value Functions	directly	Kernel extraction
		Choice			
Stakes	Alternatives	Rating	Performance Indicators		Sorting algorithms
				indirectly by learning	
Resources	Performance Criteria	Clustering	Preference modelling		Quantiles estimation

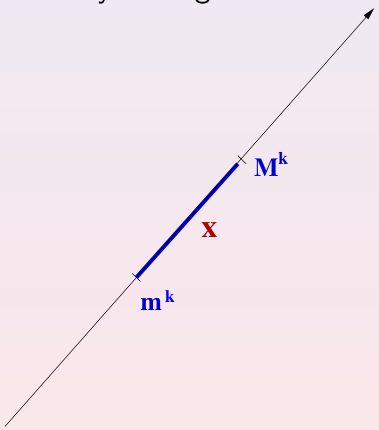
Content of the lecture Introduction Decision aiding Recommendations Bibliography

The k-Rating Problem

- A rating problem consists in a supervised partitioning of the set of alternatives into $k = 2, \ldots$. ordred categories.
- Usually, a rating procedure is designed to deal with an absolute 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 type of such norms are usually provided :
  - Delimiting (min-max) evaluation profiles ;
  - Central representatives.

Rating with delimiting norms

Rating category K is delimited by an interval $[m^k; M^k[$ on a performance measurement scale ; x is a measured performance. We may distinguish three rating situations :



1. $x < m^k$ (and $x < M^k$)
   The performance x is lower than category K ;
2. $x \geqslant m^k$ and $x < M^k$
   The performance x belongs to category K ;
3. ($x \geqslant m^k$ and) $x \geqslant M^k$
   The performance x is higher than category K.

If the relation $<$ is the dual of $\geqslant$, it will be sufficient to check that $x \geqslant m_k$ as well as $x \not\geqslant M_k$ are true for x to be a member of K.

Content of the lecture Introduction Decision aiding Recommendations Bibliography

Characterising the category K membership

Let $m^k = (m_1^k, m_2^k, \ldots, m_p^k)$ denote the lower limits and $M^k = (M_1^k, M_2^k, \ldots, M_p^k)$ the corresponding upper limits of category K on the criteria.

Proposition

That object x belongs to category K may be characterised as follows :

$$\boxed{r(x \in K) \;=\; \min\big(r(x \succsim m^k), -r(x \succsim M^k)\big)}$$

Relational Clustering

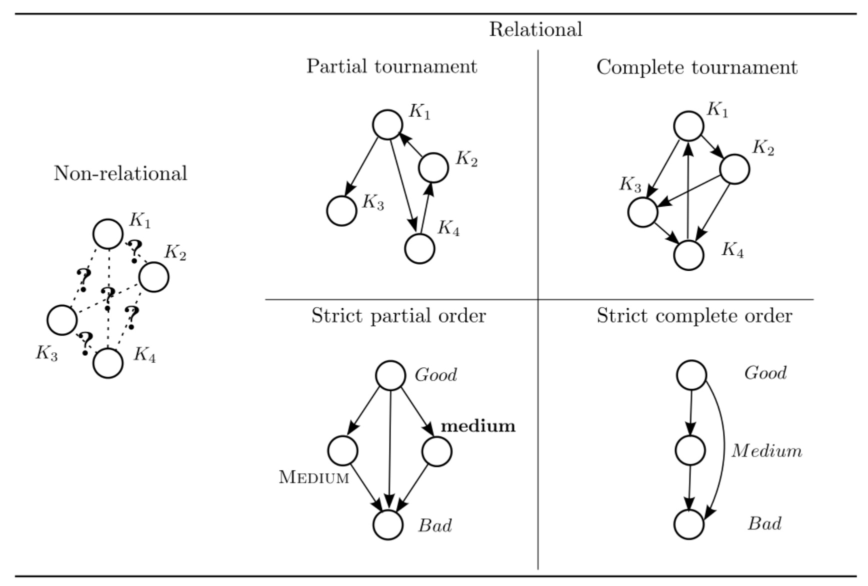
Timeline : $\longrightarrow$

Situating the problem	Formulating the problem: Decision Objects	Formulating the problem: Decision Result	Selecting the evaluation model: Evaluation model	Selecting the evaluation model: Tuning the parameters	Constructing recommendations
Actors	Objectives	Ranking	Value Functions	directly	Clique extraction
		Choice			
Stakes	Alternatives	Rating	Performance Indicators		Sorting heuristics
Resources	Performance Criteria	Clustering	Preference modelling	indirectly by learning	Quantiles estimation

The Clustering Problem

- 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.

Classification of clustering approaches



Algorithmic Approach

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

  $\rightarrow$ **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.

Algorithic Decision Theory Software Resources

- The DIGRAPH3 Linux & MacOS software collection provides practical tools for practical *Algorithmic Decision Theory* Applications.
- Download options :
  1. By using a subversion check out :
     ...$ svn co
     https ://leopold-loewenheim.uni.lu/svn/repos/Digraph3
  2. By using a github clone :
     ...$ git clone https ://github.com/rbisdorff/Digraph3
  3. Or a sourceforge clone :
     ...$ git clone https ://git.code.sf.net/p/digraph3/code
     Digraph3
- Tutorials and Reference Manual :
  https://digraph3.readthedocs.io/en/latest/

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.

(See https://leopold-loewenheim.uni.lu/bisdorff/publications.html)