Algorithmic Decision Theory

Raymond Bisdorff

Computer Science and Communications RU
Faculty of Sciences, Technology & Communication
University of Luxembourg
http://charles-sanders-peirce.uni.lu/bisdorff/

ICOCBA'2012, Kolkata December 2012

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. The GDRI AlgoDec
   - History
   - Members
   - Online Resources
2. Introduction to Algorithmic Decision Theory
   - Formulating decision problems
   - Identifying a decision problematique
   - Selecting the evaluation model
3. Constructing decision recommendations
   - Selecting and ranking
   - K-sorting
   - Clustering

Algorithmic Decision Theory

- From 2007 to 2011 the COST Action IC0602 *Algorithmic Decision Theory* (coordinated by Alexis Tsoukiàs) put together 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 Dcision Making
  - Decision Theoretic Artificial Intelligence
  - Preferences in Reasoning and Decision
  - Knowledge extraction and Learning

The GDRI ALGODEC

- The *Groupement de Recherche International* GDRI ALGODEC follows from the COST Action IC0602 *Algorithmic Decision Theory* and federates a number of 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 established in 2011 around the following institutions :

DIMACS - Rutgers University (Rutgers, USA)

CNRS - Centre National de la Recherche Scientifique (FR)

LAMSADE - Université Paris-Dauphine (Paris, FR)

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

CRIL - Université d'Artois (Lens, FR)

FNRS - Fonds National de la Recherche Scientifique (BE)

MATHRO - Université de Mons (BE)

SMG - Université Libre de Bruxelles (BE)

FNR - Fonds National de la Recherche (LU)

ILIAS - University of Luxembourg (LU)

DEIO - Universidad Rey Juan Carlos (Madrid, ES)

ALGODEC Coordination

The activities of the GDRI ALGODEC are steered by the following Committee :

Dr. Denis BOUYSSOU, **Coordinator**, LAMSADE (Paris)

Prof. Raymond BISDORFF, ILIAS (Luxembourg)

Prof. Yves DE SMET, SMG (Bruxelles)

Prof. Pierre MARQUIS, CRIL (Lens, FR)

Prof. Patrice PERNY, LIP6 (Paris)

Prof. Marc PIRLOT, MATHRO (Mons, BE)

Prof. David RIOS INSUA, DEIO (Madrid)

Prof. Fred ROBERTS, DIMACS (Rutgers USA)

D. Bouyssou

ALGODEC Activities

- EURO working group on Multiple Criteria Decision Aid
- The 1st and 2nd International Conference on Algorithmic Decision Theory
- The DECISION DECK project
- EURO working group on Preference Handling
- The DIMACS Special Focus on Algorithmic Decision Theory
- The 4th International Workshop on Computational Social Choice
- Smart Cities Workshop
- Workshop on Policy Analytics
- DA2PL'2012 workshop on Multiple Criteria Decision Aid and Preference Learning

GDRI ALGODEC Online Resources

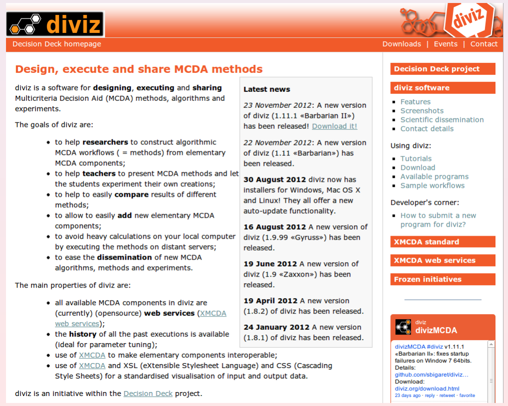
Tutorials and course materials on `http://www.algodec.org`.

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.

DECISION DECK Online Resources

XMCDA data encoding resources and webservices on `http://www.decision-deck.org`.

The DECISION DECK project aims at collaboratively developing Open Source software tools implementing MultiCriteria Decision Aid (MCDA) techniques which are meant to support complex decision aid processes and are interoperable in order to create a coherent ecosystem.



1. The GDRI ALGODEC
History
Members
Online Resources

2. Introduction to Algorithmic Decision Theory
Formulating decision problems
Identifying a decision problematique
Selecting the evaluation model

3. Constructing decision recommendations
Selecting and ranking
K-sorting
Clustering

Characterizing Decision Problems

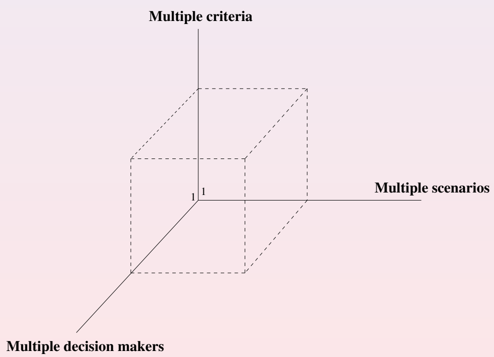
A decision problem will be a tuple $\mathcal{P} = (D, A, O, F, \Omega)$ where

1. D is a group of $d = 1, \ldots$ decision makers ;
2. A is a set of $n = 2, \ldots$ decision actions or 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 (to be maximised or minimised) with respect to decision objective $obj \in O$;
5. Ω is a set of $\omega = 1, \ldots$ potential states of the world or context scenarios.

Types of Decision Problems

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 approach

Situating the problem	Formulating the problem		Selecting the evaluation model		Constructing the final recommendation
	Decision Objects	Decision Problem	The model	Tunning the model	
Actors	Alternatives	Choice	Value Functions	directly	
Stakes		Ranking			Method
		Sorting			
Resources	Criteria		Outranking Relations	indirectly	

[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 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).

Decision Problematiques

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

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

Modelling outranking relations

X : Finite set of n alternatives

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

1. there is a (weighted) majority of voters or criteria supporting that x is at least as good as y, and
2. no veto or 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 voters or criteria supporting that x is not at least as good as y, and
2. no counter-veto or 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 $[-1, 1]$ characterizing 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 disjonction via the *max* operator,

$r(x \succsim y \wedge x' \succsim y') = \min(r(x \succsim y), r(x' \succsim y'))$
logical conjonction via the *min* operator.

1. The GDRI AlgoDec
   - History
   - Members
   - Online Resources
2. Introduction to Algorithmic Decision Theory
   - Formulating decision problems
   - Identifying a decision problematique
   - Selecting the evaluation model
3. Constructing decision recommendations
   - Selecting and ranking
   - K-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 & Bouyssou (1993), Bisdorff, Roubens & Meyer (2008).

Content of the lecture GDRI AlgoDec Algorithmic Decision Theory Decision aiding Bibliography

Useful choice qualifications

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

- Y is said to be outranking (resp. outranked) iff $x \notin X \Rightarrow \exists y \in Y : r(x \succsim y) > 0)$.
- Y is said to be independent iff for all $x \neq y$ in Y we have X $r(x \succsim y) \leqslant 0)$.
- 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 $\geqslant 1$.

Content of the lecture GDRI AlgoDec Algorithmic Decision Theory Decision aiding Bibliography

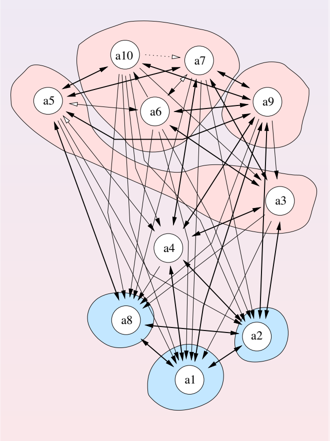
Translating BCR principles into choice qualifications

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.



Theorem (BCR Decisiveness, Bisdorff et al. 2008)

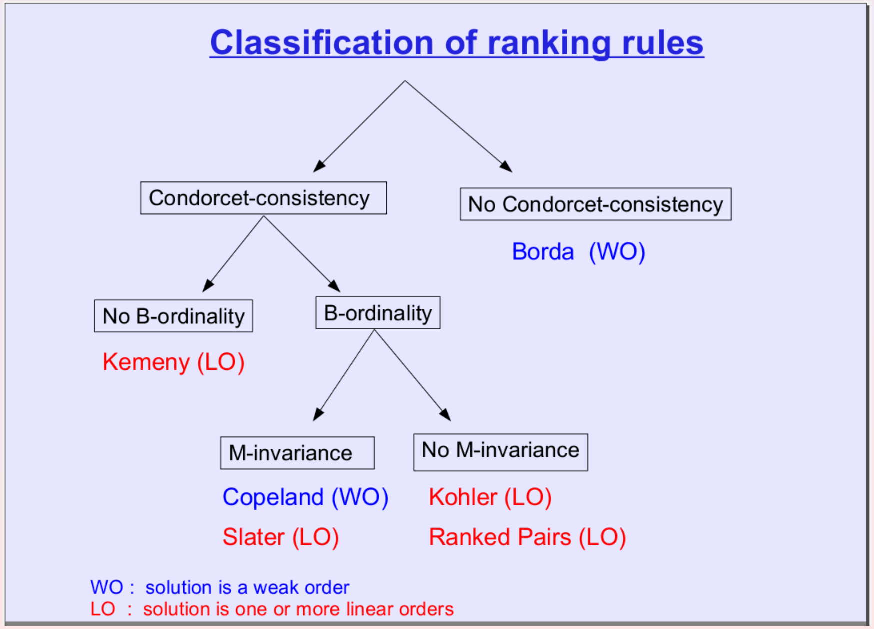
Any bipolar strict outranking digraph contains at least one outranking and one outranked hyperkernel.

Content of the lecture GDRI AlgoDec Algorithmic Decision Theory Decision aiding Bibliography

The Linear Ranking Problematique

- 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)

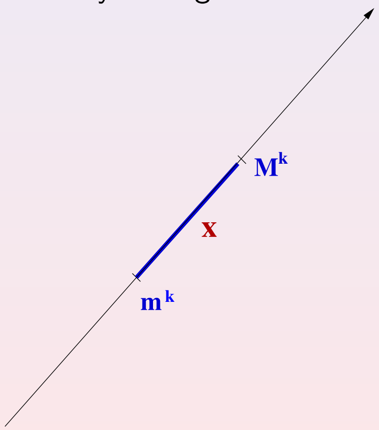


The sorting problematique

- A sorting problem consists in a supervised partitioning of the set of alternatives into $k = 2, ...$ ordred sorts or categories.
- Usually, a sorting 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.

Sorting with delimiting norms

Sort 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 sorting 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.

Notations

- $X = \{x, y, z, ...\}$ is a finite set of objects to be sorted.
- $F = \{1, ..., n\}$ is a finite and coherent family of performance criteria.
- For each criterion i in F, the objects are evaluated on a real performance scale $[0; M_i]$,

  supporting an indifference threshold q_i

  and a preference threshold p_i such that $0 \leqslant q_i < p_i \leqslant M_i$.
- The performance of object x on criterion i is denoted x_i.
- Each criterion i in F carries a rational significance w_i such that $0 < w_i < 1.0$ and $\sum_{i \in F} w_i = 1.0$.

Performing marginally *at least as good as*

Each criterion i is characterising a double threshold order $\geqslant_i$ on X in the following way :

$$r(x \geqslant_i y) = \begin{cases} +1 & \text{if} \quad x_i + q_i \geqslant y_i \\ -1 & \text{if} \quad x_i + p_i \leqslant y_i \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

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

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

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

Performing globally *at least as good as*

Each criterion i contributes the significance w_i of his "*at least as good as*" characterisation $r(\geqslant_i)$ to the global characterisation $r(\succsim)$ in the following way :

$$\boxed{r(x \succsim y) \;=\; \sum_{i \in F} \left[\, w_i \cdot r(x \geqslant_i y) \,\right]} \qquad (2)$$

$r > 0$ signifies x is *globally performing at least as good as* y,

$r < 0$ signifies that x is *not globally performing at least as good as* y,

$r = 0$ signifies that it is *unclear* whether x is globally performing at least as good as y.

Performing marginally and globally *less than*

Each criterion i is characterising a double threshold order $<_i$ (*less than*) on X in the following way :

$$r(x <_i y) \quad = \quad \begin{cases} +1 & \text{if} \quad x_i + p_i \leqslant y_i \\ -1 & \text{if} \quad x_i + q_i \geqslant y_i \\ 0 & \text{otherwise.} \end{cases} \qquad (3)$$

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

$$\boxed{r(x \prec y) = \sum_{i \in F} \left[w_i \cdot r(x <_i y) \right]} \qquad (4)$$

Proposition (Bisdorff 2012)

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

Characterising the category K membership

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

Proposition

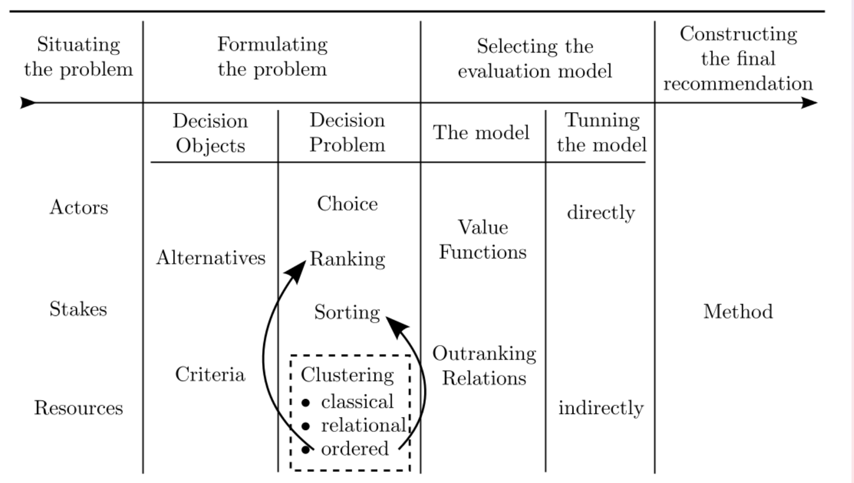
That object x belongs to category K may be characterised as follows :

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

The clustering problematique

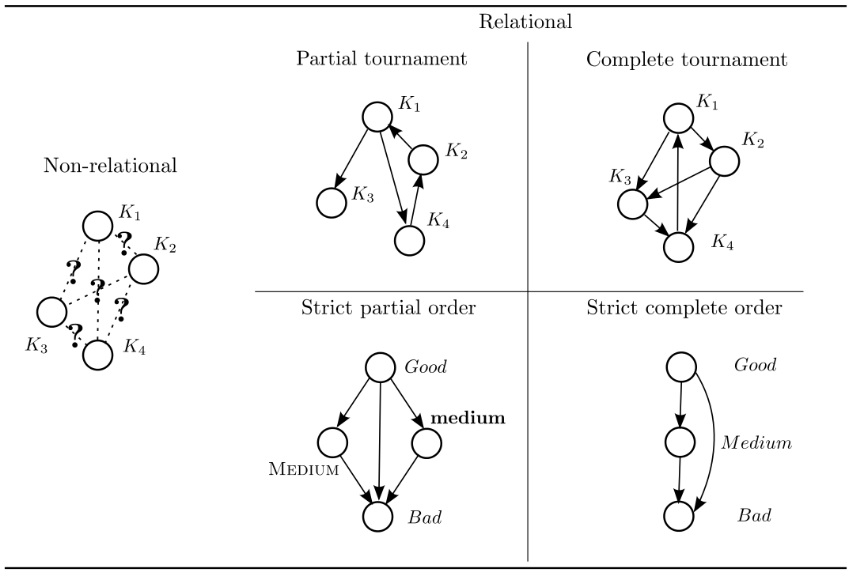
- Clustering is an unsupervised learning method that groups a set of objects into clusters.
- Properties :
  - Unknown number of clusters ;
  - Unknown characteristics of clusters ;
  - Only the relations between objects are used ;
  - no relation to external categories are used.
- Usually used in exploratory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

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.

Bibliography I

- F. Rossi and A. Tsoukiàs, *Algorithmic Decision Theory*. Proceedings of the 1st International Conference ADT2009, LNAI 5783, Springer Verlag, Berlin, 460 pages, 2009.
- D. Bouyssou, Th. Marchant, M. Pirlot, A. Tsoukiàs and Ph. Vincke, *Evaluation and Decision Models : stepping stones for the analyst*. Springer Verlag, Berlin, 2006.
- D. Bouyssou, Th. Marchant, P. Perny, M. Pirlot, A. Tsoukiàs and Ph. Vincke, *Evaluation and Decision Models : a critical perspective*. Kluwer Academic, Dordrecht, 2000.
- 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

- P. Perny and A. Tsoukiàs, *Decision analysis and artificial intelligence*. European Journal of Operational Research, vol. 160(3), 300 pages, 2005.
- A. Tsoukiàs, *From Decision Theory to Decision Aiding Methodology*. European Journal of Operational Research, vol. 187 (2008) 138 - 161.
- R. Bisdorff, *On polarizing outranking relations with large performance differences*. Journal of Multi-Criteria Decision Analysis DOI : 10.1002/mcda.1472 1-20
- 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.
- 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.

Bibliography III

- 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.
- 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, 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.