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

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 Tsoukiàis) 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 Decision 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 area.

- 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 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


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.

Characterizing Decision Problems

A decision problem will be a tuple \( P = (D, A, O, F, \Omega) \) 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 \( \text{obj} \in O \);
5. \( \Omega \) 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

<table>
<thead>
<tr>
<th>Situating the problem</th>
<th>Formulating the problem</th>
<th>Selecting the evaluation model</th>
<th>Constructing the final recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actors</td>
<td>Decision Objects</td>
<td>Decision Problem</td>
<td>The model</td>
</tr>
<tr>
<td>Alternatives</td>
<td>Choice</td>
<td>Value Functions</td>
<td>directly</td>
</tr>
<tr>
<td>Stakes</td>
<td>Ranking</td>
<td>Sorting</td>
<td>Method</td>
</tr>
<tr>
<td>Resources</td>
<td>Criteria</td>
<td>Outranking Relations</td>
<td>indirectly</td>
</tr>
</tbody>
</table>

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

Modelling outranking relations

$X$ : Finite set of $n$ alternatives

$x \preceq y$ : 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.

$x \npreceq y$ : 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 criterion.

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

Epistemic truth semantics of the $r$-valuation

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

$r(x \preceq y) = +1$ means that assertion $x \preceq y$ is *certainly valid*,

$r(x \preceq y) = -1$ means that assertion $x \preceq y$ is *certainly invalid*,

$r(x \preceq y) > 0$ means that assertion $x \preceq y$ is more *valid* than invalid,

$r(x \preceq y) < 0$ means that assertion $x \preceq y$ is more *invalid* than valid,

$r(x \preceq y) = 0$ means that validity of assertion $x \preceq y$ is *indeterminate*,

$r(x \preceq y) > r(x' \preceq y')$ means that assertion $x \preceq y$ is *more valid* than assertion $x' \preceq y'$,

$r(\neg x \preceq y) = -r(x \preceq y)$ logical (strong) negation by changing sign,

$r(x \preceq y \lor x' \preceq y') = \max (r(x \preceq y), r(x' \preceq y'))$ logical disjunction via the $\max$ operator,

$r(x \preceq y \land x' \preceq y') = \min (r(x \preceq y), r(x' \preceq y'))$ logical conjunction via the $\min$ operator.
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 retainment 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.


Useful choice qualifications

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

- $Y$ is said to be outranking (resp. outranked) iff $x \not\in X \Rightarrow \exists y \in Y : r(x \succcurlyeq y) > 0$.
- $Y$ is said to be independent iff for all $x \neq y$ in $Y$ we have $X \ni x \succcurlyeq y \leq 0$.
- $Y$ is called an outranking kernel (resp. outranked kernel) iff it is an outranking (resp. outranked) and independent 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 $\geq 1$.

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 strictly more outranking than outranked choice.
- $P_5$: Maximal credibility. A BCR has maximal determinateness.

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 $M$.

The sorting problematique

- A sorting problem consists in a supervised partitioning of the set of alternatives into \( k = 2, \ldots \) ordered 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. \( x \geq m^k \) and \( x < M^k \)
   - The performance \( x \) belongs to category \( K \).
3. \( (x \geq m^k \) and \( x \geq M^k \)
   - The performance \( x \) is higher than category \( K \).

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

Notations

- \( X = \{x, y, z, \ldots\} \) is a finite set of objects to be sorted.
- \( F = \{1, \ldots, n\} \) 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 \( 0 \leq q_i < p_i \leq M_i \).
- 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 \textit{at least as good as}

Each criterion \(i\) is characterising a double threshold order \(\geq_i\) on \(X\) in the following way:

\[
    r(x \geq_i y) = \begin{cases} 
        +1 & \text{if } x_i + q_i \geq y_i \\
        -1 & \text{if } x_i + p_i \leq y_i \\
        0 & \text{otherwise.} 
    \end{cases} \tag{1}
\]

+1 signifies \(x\) is \textit{performing at least as good as} \(y\) on criterion \(i\),

\(-1\) signifies that \(x\) is \textit{not performing at least as good as} \(y\) on criterion \(i\),

\(0\) signifies that it is \textit{unclear} whether, on criterion \(i\), \(x\) is performing at least as good as \(y\).

Performing marginally and globally \textit{less than}

Each criterion \(i\) is characterising a double threshold order \(<_i\) (\textit{less than}) on \(X\) in the following way:

\[
    r(x <_i y) = \begin{cases} 
        +1 & \text{if } x_i + p_i \leq y_i \\
        -1 & \text{if } x_i + q_i \geq y_i \\
        0 & \text{otherwise.} 
    \end{cases} \tag{3}
\]

And, the \textit{global less than} relation \((<_\cdot)\) is defined as follows:

\[
    r(x < y) = \sum_{i \in F} [w_i \cdot r(x <_i y)] \tag{4}
\]

Proposition (Bisdorff 2012)

The global \textit{“less than”} relation \(<_\cdot\) is the \textit{dual} \((\not\supseteq\cdot\) of the global \textit{“at least as good as”} relation \(\supseteq\cdot\).

Performing globally \textit{at least as good as}

Each criterion \(i\) contributes the significance \(w_i\) of his \textit{“at least as good as”} characterisation \(r(\geq_i)\) to the global characterisation \(r(\geq)\) in the following way:

\[
    r(x \supseteq y) = \sum_{i \in F} [w_i \cdot r(x \geq_i y)] \tag{2}
\]

\(r > 0\) signifies \(x\) is \textit{globally performing at least as good as} \(y\),

\(r < 0\) signifies that \(x\) is \textit{not globally performing at least as good as} \(y\),

\(r = 0\) signifies that it is \textit{unclear} whether \(x\) is globally performing at least as good as \(y\).

Characterising the category \(K\) membership

Let \(m^k = (m^k_1, m^k_2, \ldots, m^k_p)\) denote the \textit{lower limits} and \(M^k = (M^k_1, M^k_2, \ldots, M^k_p)\) the corresponding \textit{upper limits} of category \(K\) on the criteria.

Proposition

\textit{That object} \(x\) \textit{belongs to category} \(K\) \textit{may be characterised as follows}:

\[
    r(x \in K) = \min \left( r(x \supseteq m^k), -r(x \supseteq M^k) \right) 
\]
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.

<table>
<thead>
<tr>
<th>Clustering decision aid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Situating the problem</td>
</tr>
<tr>
<td>Formulating the problem</td>
</tr>
<tr>
<td>Selecting the evaluation model</td>
</tr>
<tr>
<td>Constructing the final recommendation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Actors</td>
</tr>
<tr>
<td>Alternatives</td>
</tr>
<tr>
<td>Stakes</td>
</tr>
<tr>
<td>Resources</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Decision Objects</td>
</tr>
<tr>
<td>Decision Problem</td>
</tr>
<tr>
<td>The model</td>
</tr>
<tr>
<td>Tuning the model</td>
</tr>
<tr>
<td>Choice</td>
</tr>
<tr>
<td>Ranking</td>
</tr>
<tr>
<td>Sorting</td>
</tr>
<tr>
<td>Outranking Relations</td>
</tr>
<tr>
<td>Method</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Classification of clustering approaches

<table>
<thead>
<tr>
<th>Relational</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial tournament</td>
</tr>
<tr>
<td>Complete tournament</td>
</tr>
<tr>
<td>Non-relational</td>
</tr>
<tr>
<td>Strict partial order</td>
</tr>
<tr>
<td>Strict complete order</td>
</tr>
</tbody>
</table>

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


Bibliography II


Bibliography III

- 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.