Multiple Criteria Decision Aiding by Constructive Preference Learning

Roman Słowiński
Institute of Computing Science, Poznań University of Technology
Systems Research Institute, Polish Academy of Sciences

18th Conference on Computer Science and Intelligence Systems
FedCSIS 2023, Warsaw, 18 September, 2023
Plan

- Introduction – where is the challenge?
- Why Multiple Criteria Decision Aiding?
- Evolution of MCDA influenced by the Artificial Intelligence paradigm
- Robust ordinal regression for value function preference model
  - Extreme ranking analysis
  - Stochastic ordinal regression
  - Robust Ordinal Regression for hierarchy of criteria
- Robust ordinal regression for outranking relation preference model
- Robust ordinal regression for decision rule preference model
- Examples of applications
- Summary and conclusions
Introduction – where is the challenge?
Decision problem

- There is an **objective** or **objectives** to be attained
- There are **many alternative ways** for attaining the objective(s) – they constitute a **set of actions** $A$ (alternatives, solutions, objects, acts, ...)
- **Questions** with respect to set $A$:

  $P_\alpha :$ *How to choose* the best action?

  $P_\beta :$ *How to classify* actions into pre-defined decision classes?

  $P_\gamma :$ *How to order* actions from the best to the worst?
Decision problem

Chosen subset $A'$ of best actions

Rejected subset $A \setminus A'$ of actions

Partial or complete ranking of actions

Class 1

Class 2

Class $p$

Class 1 $\succ$ Class 2 $\succ \ldots \succ$ Class $p$
Coping with multiple dimensions in Decision Aiding

- Decision problems $P_{\alpha}, P_{\beta}, P_{\gamma}$ involve vector evaluations of actions coming from:
  - multiple decision makers (voters, group decision)
  - multiple evaluation criteria (multiple objectives)
  - multiple possible states of the world that imply multiple consequences of the actions (probabilities of outcomes)
Multi-dimensional decision problems

<table>
<thead>
<tr>
<th>Element of set $A$</th>
<th>Social Choice (Group Decision)</th>
<th>Multiple Criteria Decision Aiding</th>
<th>Decision under Risk and Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate</td>
<td>Candidate</td>
<td>Action</td>
<td>Act</td>
</tr>
<tr>
<td>Voter</td>
<td>Voter</td>
<td>Criterion</td>
<td>Probability of an outcome</td>
</tr>
<tr>
<td>Dominance relation</td>
<td>Dominance relation</td>
<td>Dominance relation</td>
<td>Stochastic dominance relation</td>
</tr>
</tbody>
</table>

➢ The only objective information one can draw from the statement of a multi-dimensional decision problem is the **dominance relation**
SC&GD

<table>
<thead>
<tr>
<th>Voters</th>
<th>V₁</th>
<th>V₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>b</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>c</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

V₁: b > c > a
V₂: a > b > c

MCDA

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Action</th>
<th>Time</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

DRU

<table>
<thead>
<tr>
<th>Probability of gain</th>
<th>Act</th>
<th>Gain &gt; G₁</th>
<th>Gain &gt; G₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0.7</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>1.0</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>0.8</td>
<td>0.4</td>
<td></td>
</tr>
</tbody>
</table>

G₁ < G₂
Enriching dominance relation – preference modeling/learning

- Dominance relation is too poor – it leaves many actions non-comparable
- One can „enrich” the dominance relation, using preference information elicited from the DM
- Preference information is an input to learn/build a preference model that aggregates the vector evaluations of actions
- The preference model induces a preference relation in set $A$, richer than the dominance relation (the elements of $A$ become more comparable)
- A proper exploitation of the preference relation in $A$ leads to a recommendation in terms of choice, classification or ranking
- In this talk, we will consider multiple criteria decision aiding
Aggregation of multiple criteria evaluations – preference models

- Three families of **preference modeling (aggregation) methods**:
  - **Multiple Attribute Utility Theory (MAUT)** using a value function,
    \[ U(a) = \sum_{i=1}^{n} w_i g_i(a), \quad U(a) = \sum_{i=1}^{n} u_i[g_i(a)] , \]
    Choquet/Sugeno integral
  - **Outranking methods** using an outranking relation \( S = \{ \sim, \succsim, \succ, \preceq, \precsim, \preceq \} \)
    \[ a \succsim b = \text{“}a \text{ is at least as good as } b\text{”} \]
  - **Decision rule approach** using a set of decision rules
    \( \text{e.g., “If } g_i(a) \succ r_i \text{ & } g_j(a) \succ r_j \text{ & } \ldots \text{ } g_h(a) \succ r_h, \text{ then } a \rightarrow \text{Class } t \text{ or higher”} \)
    \( \text{“If } g_i(a) \succsim^h g_i(b) \text{ & } g_j(a) \succsim^h g_j(b) \text{ & } \ldots \text{ } g_p(a) \succsim^h g_p(b), \text{ then } a \succsim b \)

- Decision rule model is the most general of all three

Multiple-criteria approach over mono-criterion approach

- Operations Research was originally focused on **mono-criterion optimization** – mathematical programming, MAUT (utility function)

- A decision maker (DM) seldom has a single clear criterion in mind. Usually, there is **no common unit** for all scales of criteria, which are rather heterogeneous, so it may be very difficult to define a priori a unique criterion able to take into account **all relevant points of view**

- By making a family of criteria explicit, the **multiple-criteria approach** preserves the original concrete meaning of the corresponding evaluations for each actor, without resorting to any fabricated conversion (**the nightmare of composite indicators**)
Experiments show systematic violation of expected utility hypotheses
Von Neumann-Morgenstern utility function – certainty effect

**Expected utility function is linear in the probabilities**

\[ u_i(x_i^3) = p_i u_i(x_i^1) + (1-p_i) u_i(x_i^2) \]

\[ U(48) > 0.33 \times U(55) + 0.66 \times U(48) + 0.01 \times 0 \quad \Rightarrow \quad 0.34 \times U(48) \geq 0.33 \times U(55) \]

\[ 0.33 \times U(55) + 0.67 \times 0 > 0.34 \times U(48) + 0.66 \times 0 \quad \Rightarrow \quad 0.34 \times U(48) \leq 0.33 \times U(55) \]

Kahneman & Tversky: people tend to overvalue a sure thing
Multiple Attribute Utility Theory vs. Multiple Criteria Decision Aiding

Bernard Roy, Université Paris Dauphine (1934–2017)
Main sources of imperfect knowledge and ill determination (Roy 1985)

1. Roy’s staring hypothesis was that realistic decision aiding takes place in the context of imperfect knowledge and ill determination.

2. The decision aiding process is carried out in a real life context that may not correspond exactly to the model on which the decision aiding is based (the map is not the territory).

3. The system of values used for evaluating the feasibility and relative interest of diverse potential actions is usually fuzzy, incomplete and influenceable.

4. Hesitation of the DM, instability of their preferences, absence of some hardly expressible criteria in the considered family make that people in their judgments violate dominance.

5. Preference information is inconsistent, vague and ambiguous.
Weak points of the aggregation by utility (value) function (MAUT)

- Utility function distinguishes only 2 possible relations between actions:
  
  **preference** relation: \( a \succ b \iff U(a) > U(b) \)
  
  **indifference** relation: \( a \sim b \iff U(a) = U(b) \)

- \( \succ \) is asymmetric (antisymmetric and irreflexive) and transitive

- \( \sim \) is symmetric, reflexive and transitive

- Transitivity of indifference is troublesome, e.g.

  - In consequence, a non-zero **indifference threshold** \( q_i \) is necessary

  - An immediate transition from indifference to preference is unrealistic, so a **preference threshold** \( p_i \geq q_i \) and a **weak preference relation** \( \succ \) are desirable

- Another realistic situation which is not modelled by \( U \) is incomparability, so a good model should include also an **incomparability relation** "?"
Four basic preference relations and an outranking relation $S$

- Four basic preference relations are: $\{\sim, \succ, \succsim, ?\}$

- Outranking relation $S$ groups three basic preference relations:
  
  $S = \{\sim, \succ, \succsim\}$ – reflexive and non-transitive
  
  $aSb$ means: "action $a$ is at least as good as action $b$"

- For each couple $a,b \in A$:
  
  $aSb \land \text{non } bSa \iff a\succ b \lor a\succ b$
  
  $aSb \land bSa \iff a\sim b$
  
  $\text{non } aSb \land \text{non } bSa \iff a?b$
The evolution of MCDA towards AI

- Aggregation of vector evaluations, i.e., **preference modeling**:
  - till early 80’s: „**model-centric**”
    (model first, then preference info in terms of model parameters)
  - since 80’s: more and more „**human-centric**”
    (PC allowed human-computer interaction – „trial-and-error”)
  - in XXI century: „**knowledge driven**”
    (more data about human choices;
    holistic preference information first, then model building;
    explanation of past decisions, and prediction of future decisions;
    AI – model and human learn in the loop of interaction)
Elicitation of preference information by the Decision Maker (DM)

- Direct or indirect?
- Direct elicitation of numerical values of model parameters by DMs demands much of their cognitive effort


**Value function model**
- substitution rates or shapes of marginal value functions

**Outranking model**
- weights & discrimination thresholds

\[
U(a) = \sum_{i=1}^{n} w_i g_i(a) \quad \text{or} \quad \sum_{i=1}^{n} u_i [g_i(a)]
\]

\[
aSb \Leftrightarrow C(a, b) = \sum_{i=1}^{n} C_i(a, b) \geq \lambda \\
\text{and } g_i(b) - g_i(a) \leq \nu_i \quad \text{for all } i
\]
Elicitation of preference information by the Decision Maker (DM)

- **Indirect** elicitation: through holistic judgments, i.e., decision examples

- Decision aiding based on decision examples is gaining importance because:
  - Decision examples are relatively "easy" preference information
  - Decisions can also be observed without active participation of DMs
  - Psychologists confirm that **DMs are more confident exercising their decisions than explaining them** (J.G. March 1978; P. Slovic 1977)

- Related paradigms:
  - **Revealed preference theory** in economics (P. Samuelson 1938), is a method of analyzing choices made by individuals: preferences of consumers can be revealed by their purchasing habits
  - **Learning from examples** in AI/ML (knowledge discovery)

- Conclusion: **indirect elicitation of preferences is more user-friendly**
Indirect elicitation of preference information by the DM

\[ \text{[TIME}=24, \text{COST}=56, \text{RISK}=75] \quad \succ \quad \text{[TIME}=28, \text{COST}=67, \text{RISK}=25] \]

\[ \text{[MATH}=18, \text{PHYS}=16, \text{LIT}=15] \quad \Rightarrow \quad \text{Class} \; \text{“MEDIUM”} \]
\[ \text{[MATH}=17, \text{PHYS}=16, \text{LIT}=18] \quad \Rightarrow \quad \text{Class} \; \text{“GOOD”} \]

A is preferred to Z more than C is preferred to K

Action F should be among 5% of the best ones
Ordinal regression paradigm
Ordinal regression paradigm (UTA method)

- Ordinal regression paradigm emphasizes the discovery of intentions expressed through decision examples

---

UTA additive preference model

The scale of $u_i$ is a conjoint interval scale whatever the scale of $g_i$

? can be found by LP

$U(x) = \sum_{i=1}^{n} u_i[g_i(x)]:$ the value of action $x$

having evaluations $g_i(x), i = 1, \ldots, n$

Criteria are supposed to be independent with respect to preferences
Example

- Ranking of countries wrt digital economy (quality of information and technology infrastructure) (Economist Intelligence Unit in 2010)
Value function reproducing pairwise comparisons is not unique

Compatible value function ranks all countries while respecting the preference information. Another compatible value function may rank the countries otherwise.

The two rankings are substantially different, although both reproduce the same preference information.
Robust Ordinal Regression for value function preference model
Non-univocal representation - Robust Ordinal Regression - UTA\textsuperscript{GMS}

\[ U(a) = \sum_{i=1}^{n} u_i[g_i(a)] \]

The possible preference relation: for all alternatives \( x, y \in A \),

\[ x \geq_P y \iff U(x) \geq U(y) \text{ for at least one compatible value function} \]

(complete and negatively transitive)

The necessary preference relation: for all alternatives \( x, y \in A \),

\[ x \geq_N y \iff U(x) \geq U(y) \text{ for all compatible value functions} \]

(partial preorder)

When there is no preference information: necessary relation = dominance relation

\[ x \geq_N y \Rightarrow x \geq_P y, \]

i.e., \( \geq_N \subseteq \geq_P \)

\[ x \geq_N y \text{ or } y \geq_P x \]

for all \( x, y \in A \)

ROR – possible and necessary preference relations
Non-univocal representation - Robust Ordinal Regression - UTA$^{GMS}$

\[ U(a) = \sum_{i=1}^{n} u_i[g_i(a)] \]

preference information:

- \( x \succeq y \)
- \( z \succeq w \)
- \( y \succeq v \)
- \( u \succeq t \)
- \( z \succeq u \)
- \( u \succeq z \)

necessary ranking
(partial preorder)
Non-univocal representation - Robust Ordinal Regression - UTA\textsuperscript{GMS}

\[ U(a) = \sum_{i=1}^{n} u_i[g_i(a)] \]

additional preference information

\[ x \succeq y \]
\[ z \succeq w \]
\[ y \succeq v \]
\[ u \succeq t \]
\[ z \succeq u \]
\[ u \succeq z \]
\[ x \succeq w \]

necessary ranking enriched
Recommendation in terms of a necessary ranking - $\text{UTA}^{\text{GMS}}$

- **Necessary preference relation** in the set of countries, obtained by all additive value functions compatible with preference information.
Robust Ordinal Regression as a constructive learning

- Robust Ordinal Regression works in a loop with incremental elicitation of preferences → constructive learning
- Results are robust, because they take into account partial preference information

Checking for the existence of a compatible value function

$$\varepsilon^* = \max \varepsilon, \text{ subject to:}$$

$$U(a^*) \geq U(b^*) + \varepsilon \quad \text{if} \quad a^* \succ b^*$$

$$U(a^*) = U(b^*) \quad \text{if} \quad a^* \sim b^*$$

$$u_i(x_i^k) - u_i(x_i^{k-1}) \geq 0, \quad i = 1, \ldots, n, \quad k = 1, \ldots, m_i(A^R)$$

$$u_i(x_i^0) = 0, \quad i = 1, \ldots, n$$

$$\sum_{i=1}^{n} u_i(x_i^{m_i}) = 1$$

If $E^{A^R}$ is feasible and $\varepsilon^* > 0$, then there exists at least one value function compatible with the preference information
Calculating necessary and possible preference relations

- For all pairs of actions $a, b \in A$, their performances on criteria $g_i(a), g_i(b)$ add to $m_i(A^R)$ characteristic points of marginal value function $u_i, i=1,...,n$; then $E^A R$ becomes $E(a,b)$

- Consider constraints:

$$
\begin{align*}
U(b) &\geq U(a) + \varepsilon & E^N(a,b) \\
E(a,b) &
\end{align*}
\quad
\begin{align*}
U(a) &\geq U(b) & E^P(a,b)
\end{align*}
$$

- The necessary and the possible preference relations (LP problems):

$$
\begin{align*}
a \succeq_N b &\iff E^N(a,b) \text{ infeasible or } \varepsilon^N(a,b) = \max \varepsilon, \text{ s.t. } E^N(a,b) \text{ is } \leq 0 \\
a \succeq_P b &\iff E^P(a,b) \text{ feasible and } \varepsilon^P(a,b) = \max \varepsilon, \text{ s.t. } E^P(a,b) \text{ is } > 0
\end{align*}
$$
When the adopted value function fails to represent preferences...

If for a given preference information there is no compatible value function, the user can:

- identify and eliminate „troublesome” pieces of preference information (Mousseau et al. 2003),
- continue to use „not completely compatible” set of value functions with an acceptable approximation error
- augment the complexity of the value function, i.e., pass from additive value function to Choquet integral or augmented additive value function taking into account interactions between criteria

Extreme ranking analysis
Extreme ranking analysis

- Collate each action with all the remaining actions jointly
- Compute the highest and the lowest ranks and scores

\[ P^*_a = f_{\text{max}}^+ + 1 \quad \text{the best case} \]
\[ P^*_a = |A| - f_{\text{min}}^+ \quad \text{the worst case} \]

Extreme ranking analysis

- **Narrow** ranges (Bulgaria) vs. **wide** ranges (UK)
- Interactive specification of **new pairwise comparisons**, e.g., (UK, Ireland), (Poland, Slovakia)
- Choice of the **best actions**, e.g., $\text{BEST} = \{a \in A : P^*(a) = 1\}$

### Preference Information

<table>
<thead>
<tr>
<th></th>
<th>Sweden</th>
<th>Denmark</th>
<th>UK</th>
<th>Norway</th>
<th>France</th>
<th>Malta</th>
<th>Bulgaria</th>
<th>Netherland</th>
<th>Poland</th>
<th>Turkey</th>
<th>Russia</th>
<th>Kazakh.</th>
<th>Ukraine</th>
<th>Azerbaj.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>1</td>
<td></td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>Denmark</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>12</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>Norway</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>1</td>
<td>12</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>4</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>4</td>
<td>14</td>
<td>16</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malta</td>
<td>4</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>4</td>
<td>14</td>
<td>16</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulgaria</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>14</td>
<td>16</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherland</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>14</td>
<td>16</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>10</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>10</td>
<td>10</td>
<td>13</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>10</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Russia</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kazakh.</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ukraine</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Azerbaj.</td>
<td>16</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic ordinal regression
Stochastic Multiobjective Acceptability Analysis & ROR = SOR

- When the necessary preference relation $\succeq^N$ is poor, it leaves many pairs of alternatives incomparable, i.e., $a \succeq^P b$ and $b \succeq^P a$

- The number of compatible value functions constrained by available preference information is infinite

- One can sample these compatible value functions within the constraints and check the frequency with which:
  - $a \succ b$ – pairwise winning index $p(a,b)$,
  - $a$ gets position $i$ in the ranking – rank acceptability index $b^i_a$

- The sampling is performed using the Hit and Run algorithm (Smith 1984) (Monte Carlo simulation)

Robust Ordinal Regression for hierarchy of criteria
Multiple Criteria Hierarchy Process (MCHP)

Multiple Criteria Hierarchy Process (MCHP) – main idea

➢ We wish to consider preference relation $\succeq_r$ in each node of the hierarchy tree, e.g.:

- $a \succeq_{(2)} b \iff U_{(2)}(a) \geq U_{(2)}(b)$
- $c \succeq_{(1,3)} d \iff U_{(1,3)}(c) \geq U_{(1,3)}(d)$
- $e \succeq_{(2,1)} f \iff U_{(2,1)}(e) \geq U_{(2,1)}(f)$

Indirect preference information in particular nodes of the tree:

- **Pairwise comparison:** \( a \) is at least as good as \( b \) on criterion \( G_r \)

\[
a \succeq_r b \iff U_r(a) \geq U_r(b)
\]

- **Intensity of preference:** considering criterion \( G_r \) or \( g_t \),

\( a \) is preferred to \( b \) at least as much as \( c \) is preferred to \( d \)

\[
(a, b) \succeq_r^*(c, d) \iff U_r(a) - U_r(b) \geq U_r(c) - U_r(d)
\]

\[
(a, b) \succeq_t^*(c, d) \iff u_t(a) - u_t(b) \geq u_t(c) - u_t(d)
\]
Properties of necessary and possible preference relations in node $r$

- Given two alternatives $a, b \in A$, and any non-elementary criterion $G_r$:

  (i) $a \succeq_{(r, j)}^N b$ for all $j = 1, \ldots, n(r)$ $\Rightarrow$ $a \succeq_r^N b$

  (ii) $a \succeq_{(r, j)}^N b$ for all $j = 1, \ldots, n(r)$, $j \neq w$, and $a \succeq_{(r, w)}^P b$ $\Rightarrow$ $a \succeq_r^P b$

  $\neg(a \succeq_{(r, j)}^P b)$ for all $j = 1, \ldots, n(r)$ $\Rightarrow$ $\neg(a \succeq_r^P b)$

  (iii) $a \succeq_r^P b$ $\Rightarrow$ $a \succeq_{(r, j)}^P b$ for at least one $j \in \{1, \ldots, n(r)\}$

- **Remark**: hierarchical properties are expressed in terms of preference
  - necessary (i)
  - necessary & possible (ii)
  - possible (iii)
Other developments in MCHP for value function and ROR:

- Choquet integral value function

- Choquet integral value function and Stochastic Ordinal Regression

- MCHP for sorting problems with additive value functions

S. Angilella, S. Corrente, S. Greco, R. Słowiński: Robust Ordinal Regression and Stochastic Multiobjective Acceptability Analysis in Multiple Criteria Hierarchy Process for the Choquet integral preference model. OMEGA, 63 (2016) 154-169
Robust Ordinal Regression for outranking relation preference model
Robust Ordinal Regression approach for outranking methods

- **Concordance test**: checks if the coalition of criteria concordant with the hypothesis \( aSb \) is strong enough:

\[
C(a,b) = \frac{\sum_{i=1}^{n} w_i C_i(a,b)}{\sum_{i=1}^{n} w_i} \quad a, b \in A, \ w_i \text{ are weights of criteria}
\]

- Concordance test is **positive** if: \( C(a,b) \geq \lambda \), where \( \lambda \in [0.5, 1] \) is a cutting level (concordance threshold)

- **No compensation** between criteria because the weights are not multiplied by performances (weight \( w_i \) is a voting power of \( g_i \))
Robust Ordinal Regression approach for outranking methods

- **Discordance test:** checks if among criteria discordant with the hypothesis $aSb$ there is a strong opposition against $aSb$:
  - $g_i(b) - g_i(a) \geq v_i$ (for gain-type criterion)
  - $g_i(a) - g_i(b) \geq v_i$ (for cost-type criterion)

- **Conclusion:** $aSb$ is true if and only if $C(a,b) \geq \lambda$ and there is no criterion strongly opposed (making veto) to the hypothesis

- For each couple $(a,b) \in A \times A$, one obtains relation $S$: true (1) or false (0)

![Table and diagram](image)
Robust Ordinal Regression approach for outranking methods

- Assuming $\sum_{i=1}^{n} w_i = 1$, we have

\[ C(a,b) = \sum_{i=1}^{n} w_i C_i(a,b) = \sum_{i=1}^{n} \psi_i(a,b) \]

where $\psi_i(a,b)$ is a non-decreasing function of $g_i(a) - g_i(b)$

where $\alpha_i, \beta_i$ are, respectively, the worst and the best possible performance on criterion $g_i$, $i=1,\ldots,n$
Robust Ordinal Regression approach for outranking methods

- Preference information provided by the DM (ELECTRE$^{GKMS}$):

  \[ aSb \text{ or } aS^c b, \text{ for } a, b \in A^R \subset A \]

  \[ [q_i^*, q_i^*] \] - the range of indifference threshold allowed by the DM

  \[ [p_i^*, p_i^*] \] - the range of preference threshold allowed by the DM

Robust Ordinal Regression approach for outranking methods

- **Compatible outranking model** is a set of marginal concordance functions $\Psi_i(a,b)$, cutting levels $\lambda$, indifference $q_i$, preference $p_i$, and veto thresholds $\nu_i$, $i=1,...,n$, reproducing the DM’s preference information concerning pairs $(a,b) \in A^R \times A^R$.
Robust ordinal regression approach for outranking methods

- Ordinal regression (compatibility) constraints $E^{AR}$:

  If $aSb$ for $(a,b) \in A^R \times A^R$:
  
  $$C(a,b) = \sum_{i=1}^{n} \psi_i(a,b) \geq \lambda$$
  
  $$g_i(b) - g_i(a) + \varepsilon \leq v_i, \ i = 1, \ldots, n$$

  If $aS^c b$ for $(a,b) \in A^R \times A^R$:
  
  $$C(a,b) = \sum_{i=1}^{n} \psi_i(a,b) + \varepsilon \leq \lambda + M_0(a,b)$$
  
  $$g_i(b) - g_i(a) \geq v_i - \delta M_i(a,b), \ i = 1, \ldots, n$$
  
  $$M_i(a,b) \in \{0, 1\}, \ i = 0, 1, \ldots, n$$
  
  $$\sum_{i=0}^{n} M_i(a,b) \leq n, \ where \ \delta \ is \ a \ big \ given \ value$$

  $0.5 \leq \lambda \leq 1,$
  
  $$v_i \geq p_i^* + \varepsilon, \ if \ \lfloor p_i^*, p_i^* \rfloor \ was \ given$$
  
  $$v_i \geq g_i(b) - g_i(a) + \varepsilon, \ v_i \geq g_i(a) - g_i(b) + \varepsilon, \ if \ a \sim b \ was \ given, \ i \in \{1, \ldots, n\}$$
Robust Ordinal Regression approach for outranking methods

- Given a pair of alternatives \( a, b \in A \), \( a \) necessarily outranks \( b \):
  \[
  a S^N b \iff \epsilon^* \leq 0
  \]
  where \( \epsilon^* = \max \epsilon \)

  subject to:
  \[
  E^A^R
  \]
  \[
  \begin{aligned}
  C(a, b) &= \sum_{i=1}^{n} \psi_i(a, b) + \epsilon \leq \lambda + M_0(a, b) \\
  g_i(b) - g_i(a) &\geq v_i - \delta M_i(a, b) \\
  M_i(a, b) &\in \{0, 1\}, \quad i = 1, \ldots, n, \quad \sum_{i=0}^{n} M_i(a, b) \leq n
  \end{aligned}
  \]

- If \( \epsilon^* \leq 0 \) and constraints \( E^N(a, b) \) are infeasible, then \( a \) outranks \( b \) for all compatible outranking models (\( a S^N b \) because \( a S^{CN} b \) is not possible)
Exploitation of outranking relations $S^N$, $S^{CN}$, $S^P$, $S^{CP}$ in set $A$

- **Choice problem:**
  
  Kernel of the necessary outranking graph $S^N$

- **Ranking problem:**

  Exploitation of the necessary outranking graph including $S^N$ and $S^{CN}$ using **Net Flow Score procedure** for each alternative $x \in A$:

  $$NFS(x) = strength(x) - weakness(x)$$

  $S^N$ – positive argument, $S^{CN}$ – negative argument

  Ranking: complete preorder determined by $NFS(x)$ in $A$
Robust Ordinal Regression approach for **outranking methods**

**Necessary outranking**

<table>
<thead>
<tr>
<th>S^N</th>
<th>D</th>
<th>U</th>
<th>M</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>B</th>
<th>T</th>
<th>K</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>U</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>M</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>I</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>K</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**kernel**

![Kernel Diagram](image)

**NFS ranking**

![NFS Ranking Diagram](image)
Robust Ordinal Regression for decision rule preference model
Syntax of monotonic decision rules

**Ordinal classification**

\[
\text{if } x_{q1} \succeq r_{q1} \text{ and } x_{q2} \succeq r_{q2} \text{ and } \ldots \text{ and } x_{qp} \succeq r_{qp}, \text{ then } x \rightarrow \text{ class } t \text{ or better}
\]

\[
\text{if } x_{q1} \preceq r_{q1} \text{ and } x_{q2} \preceq r_{q2} \text{ and } \ldots \text{ and } x_{qp} \preceq r_{qp}, \text{ then } x \rightarrow \text{ class } t \text{ or worse}
\]

**Choice ranking**

\[
\text{if } (x \succ q_1 \geq h(q_1) y) \text{ and } (x \succ q_2 \geq h(q_2) y) \text{ and } \ldots \text{ and } (x \succ q_p \geq h(q_p) y), \text{ then } xSy
\]

\[
\text{if } (x \preceq q_1 \leq h(q_1) y) \text{ and } (x \preceq q_2 \leq h(q_2) y) \text{ and } \ldots \text{ and } (x \preceq q_p \leq h(q_p) y), \text{ then } xS^c y
\]

**Cardinal criteria**

\[
\text{if } x_{g1} \succeq g_1 r_{q1} \text{ and } y_{g1} \preceq g_1 r'_{q1} \text{ and } \ldots \text{ and } x_{gp} \succeq gp r_{gp} \text{ and } y_{gp} \preceq gp r'_{gp}, \text{ then } xSy
\]

\[
\text{if } x_{g1} \preceq g_1 r_{q1} \text{ and } y_{g1} \succeq g_1 r'_{q1} \text{ and } \ldots \text{ and } x_{gp} \preceq gp r_{gp} \text{ and } y_{gp} \succeq gp r'_{gp}, \text{ then } xS^c y
\]

Pair of objects \(x,y\) evaluated on criterion \(g_1\)

---

Dominance-based Rough Set Approach (DRSA)

Classes: ▲ ▼ □

Bipolarity

Lower approximation ‘at least’ class ▲
Upper approximation ‘at least’ class ▲
Boundary ‘at least’ class ▲ & ‘at most’ class □
Upper approximation ‘at most’ class □
Lower approximation ‘at most’ class □

Dominance principle (comonotonicity)

If $x$ is at least as good as $y$ with respect to relevant criteria, then $x$ should be assigned to a class not worse than $y$

Preference modeling by dominance-based decision rules

- Dominance-based „if..., then...” decision rules are the only aggregation operators that:
  - give account of most complex interactions among attributes,
  - are non-compensatory,
  - accept ordinal evaluation scales and do not convert ordinal evaluations into cardinal ones,

- Rules identify values that drive DM’s decisions – each rule is a scenario of a causal relationship between evaluations on a subset of attributes and a comprehensive judgment
Example

Sample of 8 actions submitted to evaluation of the DM

<table>
<thead>
<tr>
<th>action</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>2</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>$x_2$</td>
<td>3</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>$x_3$</td>
<td>5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>$x_4$</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>$x_5$</td>
<td>8</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>$x_6$</td>
<td>11</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>$x_7$</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>$x_8$</td>
<td>10</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>
Example

Sample of 8 actions – elicitation of preferences by the DM

<table>
<thead>
<tr>
<th>action</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>2</td>
<td>14</td>
<td>bad</td>
</tr>
<tr>
<td>$x_2$</td>
<td>3</td>
<td>12</td>
<td>bad</td>
</tr>
<tr>
<td>$x_3$</td>
<td>5</td>
<td>9</td>
<td>good</td>
</tr>
<tr>
<td>$x_4$</td>
<td>7</td>
<td>8</td>
<td>good</td>
</tr>
<tr>
<td>$x_5$</td>
<td>8</td>
<td>7</td>
<td>good</td>
</tr>
<tr>
<td>$x_6$</td>
<td>11</td>
<td>6</td>
<td>bad</td>
</tr>
<tr>
<td>$x_7$</td>
<td>9</td>
<td>10</td>
<td>bad</td>
</tr>
<tr>
<td>$x_8$</td>
<td>10</td>
<td>11</td>
<td>good</td>
</tr>
</tbody>
</table>

The actions are plotted on a 2D graph with $f_2 \rightarrow \text{min}$ on the y-axis and $f_1 \rightarrow \text{min}$ on the x-axis. Green dots represent good actions, and red squares represent bad actions.
Sample of 8 actions – dominance-based rough approximations

<table>
<thead>
<tr>
<th>action</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>2</td>
<td>14</td>
<td>bad</td>
</tr>
<tr>
<td>$x_2$</td>
<td>3</td>
<td>12</td>
<td>bad</td>
</tr>
<tr>
<td>$x_3$</td>
<td>5</td>
<td>9</td>
<td>good</td>
</tr>
<tr>
<td>$x_4$</td>
<td>7</td>
<td>8</td>
<td>good</td>
</tr>
<tr>
<td>$x_5$</td>
<td>8</td>
<td>7</td>
<td>good</td>
</tr>
<tr>
<td>$x_6$</td>
<td>11</td>
<td>6</td>
<td>bad</td>
</tr>
<tr>
<td>$x_7$</td>
<td>9</td>
<td>10</td>
<td>bad</td>
</tr>
<tr>
<td>$x_8$</td>
<td>10</td>
<td>11</td>
<td>good</td>
</tr>
</tbody>
</table>
Example

Sample of 8 actions – induction of certain decision rules

<table>
<thead>
<tr>
<th>action</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>2</td>
<td>14</td>
<td>bad</td>
</tr>
<tr>
<td>$x_2$</td>
<td>3</td>
<td>12</td>
<td>bad</td>
</tr>
<tr>
<td>$x_3$</td>
<td>5</td>
<td>9</td>
<td>good</td>
</tr>
<tr>
<td>$x_4$</td>
<td>7</td>
<td>8</td>
<td>good</td>
</tr>
<tr>
<td>$x_5$</td>
<td>8</td>
<td>7</td>
<td>good</td>
</tr>
<tr>
<td>$x_6$</td>
<td>11</td>
<td>6</td>
<td>bad</td>
</tr>
<tr>
<td>$x_7$</td>
<td>9</td>
<td>10</td>
<td>bad</td>
</tr>
<tr>
<td>$x_8$</td>
<td>10</td>
<td>11</td>
<td>good</td>
</tr>
</tbody>
</table>

$D_{\leq}$:
- $r_1$: if $f_1(x) \geq 11$, then $x$ is certainly bad
  supported by $\{x_6\}$
- $r_2$: if $f_2(x) \geq 12$, then $x$ is certainly bad
  supported by $\{x_1, x_2\}$

$D_{\geq}$:
- $r_3$: if $f_1(x) \leq 8$ & $f_2(x) \leq 9$, then $x$ is certainly good
  supported by $\{x_3, x_4, x_5\}$
Example

Sample of 8 actions – induction of possible decision rules

<table>
<thead>
<tr>
<th>action</th>
<th>( f_1 )</th>
<th>( f_2 )</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>2</td>
<td>14</td>
<td>bad</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>3</td>
<td>12</td>
<td>bad</td>
</tr>
<tr>
<td>( x_3 )</td>
<td>5</td>
<td>9</td>
<td>good</td>
</tr>
<tr>
<td>( x_4 )</td>
<td>7</td>
<td>8</td>
<td>good</td>
</tr>
<tr>
<td>( x_5 )</td>
<td>8</td>
<td>7</td>
<td>good</td>
</tr>
<tr>
<td>( x_6 )</td>
<td>11</td>
<td>6</td>
<td>bad</td>
</tr>
<tr>
<td>( x_7 )</td>
<td>9</td>
<td>10</td>
<td>bad</td>
</tr>
<tr>
<td>( x_8 )</td>
<td>10</td>
<td>11</td>
<td>good</td>
</tr>
</tbody>
</table>

\[ D \leq \]

- \( r_4 \): if \( f_1(x) \geq 9 \), then \( x \) is possibly bad
  supported by \( \{x_6, x_7, x_8\} \)

- \( r_5 \): if \( f_2(x) \geq 10 \), then \( x \) is possibly bad
  supported by \( \{x_1, x_2, x_7, x_8\} \)

\[ D \geq \]

- \( r_6 \): if \( f_1(x) \leq 10 \) & \( f_2(x) \leq 11 \), then \( x \) is possibly good
  supported by \( \{x_3, x_4, x_5, x_7, x_8\} \)
Examples of applications
Mobile Emergency Triage System

- Total pediatric population >400,000
- 55,000 patient visits in the ER per year
- 3 pediatric general surgeons (supported by emergency physicians and residents)
Triage Process

Emergency Room (ER)

<table>
<thead>
<tr>
<th>Prioritization (Triage nurse)</th>
<th>Disposition (ED Physician)</th>
<th>Examination (Specialist)</th>
<th>Hospital/Clinic</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Resuscitation Immediate</td>
<td>Consult</td>
<td></td>
<td>Observation</td>
</tr>
<tr>
<td>II Emergent ≤ 15 min.</td>
<td></td>
<td></td>
<td>Observation/Clinic</td>
</tr>
<tr>
<td>III Urgent ≤ 30 min.</td>
<td></td>
<td></td>
<td>Surgery</td>
</tr>
<tr>
<td>IV Less Urgent ≤ 1 hour</td>
<td></td>
<td></td>
<td>Discharge</td>
</tr>
<tr>
<td>V Non Urgent ≤ 2 hours</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Diagnosis and treatment

Management
MET System - scrotal pain triage

**History**
- Site of pain: Both
  - Left
  - Right
- Onset of pain: Acute
  - Gradual
- Type of pain: Constant
  - Intermittent
- Vomiting: Yes
  - No

**Physical Examination**
- Cord palpable: Abnormal
  - Normal
- Cremast. reflex: Yes
  - No
- Lie: Transverse
- Testis tenderness:
  - Entire Testis
  - Not Tender
  - Posterior
  - Tender Not Specific
  - Upper Pole
- Temperature: 38.6 °C
- Swelling: Both
  - Left
  - None
  - Right

**Tests**
- WBC/HPF: 9.0
- WBC: (no value) x 1000
- Suggested: Consult (strong)
Auto loan fraud detection using dominance-based rough set approach

- Bank data: 26,187 observations including 405 fraud events
- Accuracy of models compared:

<table>
<thead>
<tr>
<th>Model</th>
<th>Class fraud detection rate [%]</th>
<th>Class non-fraud detection rate [%]</th>
<th>G-mean [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRSA-BRE</td>
<td><strong>79.09</strong></td>
<td>82.56</td>
<td>80.81</td>
</tr>
<tr>
<td>Random Forest</td>
<td>18.19</td>
<td>99.98</td>
<td>42.62</td>
</tr>
<tr>
<td>SVM</td>
<td>13.9</td>
<td>99.94</td>
<td>37.26</td>
</tr>
</tbody>
</table>

- Examples of meaningful rules:

**#1**: if (NUMBER OF INSTALMENTS $\geq$ 60) and (CAR PRICE $\geq$ 55320) and (DOWNPAYMENT TO CAR PRICE $\leq$ 0.1) and (ANNUAL TURNOVER LAST YEAR $\geq$ 198000) and (COMPANY AGE $\leq$ 2), then fraud

**#10**: if (DOWNPAYMENT TO CAR PRICE $\leq$ 0.1) and (LEGAL FORM group = capital company) and (COMPANY AGE $\leq$ 6) and (PKD group = building), then fraud

Auto loan fraud detection using dominance-based rough set approach

- Importance of attributes in terms of attribute Bayesian confirmation:
VC-DRSA performing sequential covering adapted to missing values was applied on a set of 10 000 customers (7963 exited, 2037 loyal)

Table 2: Comparison of avg. classification accuracy (%) in 10 × 10-fold cross-validation

<table>
<thead>
<tr>
<th>%mv</th>
<th>$\epsilon$-D$_{1.5}^{mv}$</th>
<th>$\epsilon$-D$_{2}^{mv}$</th>
<th>C4.5</th>
<th>NB</th>
<th>SVM</th>
<th>RF</th>
<th>MP</th>
<th>RIPP</th>
<th>OLM</th>
<th>OSDL</th>
<th>MoNGEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>75.89</td>
<td>75.89</td>
<td>75.39</td>
<td>75.98</td>
<td>70.01</td>
<td>77.05</td>
<td>75.86</td>
<td>76.52</td>
<td>57.38</td>
<td>73.74</td>
<td>69.79</td>
</tr>
<tr>
<td>5</td>
<td>75.01</td>
<td>74.52</td>
<td>75.47</td>
<td>75.57</td>
<td>69.49</td>
<td>76.05</td>
<td>74.09</td>
<td>74.52</td>
<td>53.41</td>
<td>71.63</td>
<td>68.78</td>
</tr>
<tr>
<td>10</td>
<td>73.95</td>
<td>73.39</td>
<td>74.90</td>
<td>74.49</td>
<td>68.20</td>
<td>74.75</td>
<td>72.96</td>
<td>72.55</td>
<td>51.04</td>
<td>70.15</td>
<td>66.11</td>
</tr>
<tr>
<td>15</td>
<td>73.03</td>
<td>71.74</td>
<td>73.54</td>
<td>74.08</td>
<td>68.03</td>
<td>74.27</td>
<td>71.85</td>
<td>70.17</td>
<td>50.24</td>
<td>68.79</td>
<td>65.19</td>
</tr>
<tr>
<td>20</td>
<td>72.15</td>
<td>70.98</td>
<td>72.93</td>
<td>73.72</td>
<td>66.84</td>
<td>74.04</td>
<td>70.73</td>
<td>69.93</td>
<td>50.18</td>
<td>67.82</td>
<td>64.10</td>
</tr>
<tr>
<td>25</td>
<td>70.72</td>
<td>69.92</td>
<td>72.09</td>
<td>72.50</td>
<td>66.02</td>
<td>72.92</td>
<td>69.24</td>
<td>69.14</td>
<td>50.00</td>
<td>66.56</td>
<td>62.26</td>
</tr>
</tbody>
</table>

Table 7: Top rules induced by $\epsilon$-VC-DRSA

<table>
<thead>
<tr>
<th>ID</th>
<th>Conditions</th>
<th>Decision</th>
<th>$\epsilon$</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>98</td>
<td>CreditScore ≤ 712, Age ≥ 51, IsActiveMember ≤ 0</td>
<td>Exited = 1</td>
<td>0.005</td>
<td>264</td>
</tr>
<tr>
<td>91</td>
<td>NumOfProducts_c ≥ 3, CreditScore ≤ 789, Age ≥ 35</td>
<td>Exited = 1</td>
<td>0.003</td>
<td>221</td>
</tr>
<tr>
<td>108</td>
<td>Age ≥ 49, IsActiveMember ≤ 0, CreditScore ≤ 657, HasCrCard = 1</td>
<td>Exited = 1</td>
<td>0.005</td>
<td>172</td>
</tr>
<tr>
<td>111</td>
<td>Age ≥ 46, IsActiveMember ≤ 0, Geography = Germany, NumOfProducts_g ≤ 1, CreditScore ≤ 805</td>
<td>Exited = 1</td>
<td>0.003</td>
<td>171</td>
</tr>
<tr>
<td>97</td>
<td>Age ≥ 54, IsActiveMember ≤ 0, EstimatedSalary ≤ 123646.57</td>
<td>Exited = 1</td>
<td>0.002</td>
<td>155</td>
</tr>
</tbody>
</table>
Multiobjective Optimization

\[
\begin{bmatrix}
    f_1(x) \\
    \vdots \\
    f_n(x)
\end{bmatrix} \rightarrow \text{Min (or Max)}
\]

subject to the constraints :

\[
g_1(x) \{\leq, =, \geq\} b_1 \\
\vdots \\
g_m(x) \{\leq, =, \geq\} b_m
\]

where \( x = [x_1, \ldots, x_k] \) - vector of decision variables (continuous/integer)

\( f_j(x), j=1,\ldots,n \) - real-valued objective functions

\( g_i(x), i=1,\ldots,m \) - real-valued functions of the constraints

\( b_i, i=1,\ldots,m \) - constant RHS of the constraints
Evolutionary Multiobjective Optimization (EMO)

MOCO problems are NP-hard, \#P-hard \(\rightarrow\) intractable

Even if single-objective problem is polynomially solvable, the multiobjective problem is usually NP-hard, e.g.:

- spanning tree
- min-cost flow

(Ehrgott & Gandibleux 2000)
Multiobjective Optimization – „most preferred” solution
From preference model to ranking of solutions in a population

Preference pressure in the recombination procedure

- Mating selection, crossover and mutation in generation $t$:

parents $P_t$

$p(\cdot)$ is a permutation of $\{1, \ldots, 30\}$

offspring population $Q_t$ with 30 individuals
The NSGA-II framework

- **NSGA-II**: dominance ranking of solutions from a current population

Within the same front, order the solutions with respect to the crowding distance

---

XIMEA-DRSA: Interactive EMO driven by decision rules

S. Corrente, S. Greco, B. Matarazzo, R. Słowiński: Explainable Interactive Evolutionary Multiobjective Optimization, OMEGA, 122 (2024) 102925
Example of preference information and preference model

Sample of 8 actions – induction of certain decision rules

<table>
<thead>
<tr>
<th>action</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>2</td>
<td>14</td>
<td>bad</td>
</tr>
<tr>
<td>$x_2$</td>
<td>3</td>
<td>12</td>
<td>bad</td>
</tr>
<tr>
<td>$x_3$</td>
<td>5</td>
<td>9</td>
<td>good</td>
</tr>
<tr>
<td>$x_4$</td>
<td>7</td>
<td>8</td>
<td>good</td>
</tr>
<tr>
<td>$x_5$</td>
<td>8</td>
<td>7</td>
<td>good</td>
</tr>
<tr>
<td>$x_6$</td>
<td>11</td>
<td>6</td>
<td>bad</td>
</tr>
<tr>
<td>$x_7$</td>
<td>9</td>
<td>10</td>
<td>bad</td>
</tr>
<tr>
<td>$x_8$</td>
<td>10</td>
<td>11</td>
<td>good</td>
</tr>
</tbody>
</table>

$D_{\leq}:
\begin{align*}
  r_1: & \text{ if } f_1(x) \geq 11, \text{ then } x \text{ is certainly } \text{bad} \\
  r_2: & \text{ if } f_2(x) \geq 12, \text{ then } x \text{ is certainly } \text{bad}
\end{align*}$

$D_{\geq}:
\begin{align*}
  r_3: & \text{ if } f_1(x) \leq 8 \& f_2(x) \leq 9, \text{ then } x \text{ is certainly } \text{good}
\end{align*}$

supported by $\{x_6\}$
supported by $\{x_1, x_2\}$
supported by $\{x_3, x_4, x_5\}$
Example of preference information and preference model

Sample of 8 actions – induction of possible decision rules

<table>
<thead>
<tr>
<th>action</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>2</td>
<td>14</td>
<td>bad</td>
</tr>
<tr>
<td>$x_2$</td>
<td>3</td>
<td>12</td>
<td>bad</td>
</tr>
<tr>
<td>$x_3$</td>
<td>5</td>
<td>9</td>
<td>good</td>
</tr>
<tr>
<td>$x_4$</td>
<td>7</td>
<td>8</td>
<td>good</td>
</tr>
<tr>
<td>$x_5$</td>
<td>8</td>
<td>7</td>
<td>good</td>
</tr>
<tr>
<td>$x_6$</td>
<td>11</td>
<td>6</td>
<td>bad</td>
</tr>
<tr>
<td>$x_7$</td>
<td>9</td>
<td>10</td>
<td>bad</td>
</tr>
<tr>
<td>$x_8$</td>
<td>10</td>
<td>11</td>
<td>good</td>
</tr>
</tbody>
</table>

$r_4$: if $f_1(x) \geq 9$, then $x$ is possibly bad

$r_5$: if $f_2(x) \geq 10$, then $x$ is possibly bad

$r_6$: if $f_1(x) \leq 10$ & $f_2(x) \leq 11$, then $x$ is possibly good

$D \geq$

$D \leq$

supported by $\{x_6, x_7, x_8\}$

supported by $\{x_1, x_2, x_7, x_8\}$

supported by $\{x_3, x_4, x_5, x_7, x_8\}$
**XIMEA-DRSA:** Interactive EMO driven by decision rules

1. Assign solutions to ordered non-dominated fronts $NDF_1, \ldots, NDF_q, \ldots$

2. Inside the same non-dominated front:
   a) **Calculate** for each solution $x$ the score
      
      $$score(x) = \sum_{r \in D_{\geq}(x)} e^{-\gamma(t-age(r))} - \sum_{r \in D_{\leq}(x)} e^{-\gamma(t-age(r))}$$
      
      where $D_{\geq}(x)$ - the set of rules of type $D_{\geq}$ matching $x$ (good rules),
      $D_{\leq}(x)$ - the set of rules of type $D_{\leq}$ matching $x$ (bad rules)
   
   b) **Order** solutions in each $NDF_q$ from the highest to the lowest $score(x)$

$t$ – iteration, $r$ – rule

$age(r)$ – the iteration in which rule $r$ was born

$\gamma > 0$ – coefficient of the aging speed

$$NDF_q \rightarrow \begin{cases} x^1 \\ \vdots \\ x^l \end{cases} \text{ such that:} \\
\quad x^1 \cup \ldots \cup x^l = NDF_q \\
\quad score(x^1) > \ldots > score(x^l)$$

S. Corrente, S. Greco, B. Matarazzo, R. Słowiński: Explainable Interactive Evolutionary Multiobjective Optimization, *OMEGA*, 122 (2024) 102925
From preference model to ranking of solutions in a population

Preference pressure in the recombination procedure

- Mating selection, crossover and mutation in generation $t$:

Front 1

Parents $P_t$

$p(\cdot)$ is a permutation of $\{1, ..., 30\}$

Front 2

Front $p$

Tournament of pairs

$\Rightarrow$

choose better of 2 individuals

$\Rightarrow$

choose better of 2 individuals

crossover

mutation

offspring population $Q_t$

with 30 individuals
From preference model to ranking of solutions in a population

- Selection of new population $P_{t+1}$:

![Diagram showing the selection process of a new population $P_{t+1}$ from a preference model. The diagram illustrates sorting of 60 individuals into NDFs, with each NDF further divided into $F_1$, $F_2$, and $F_3$. The process involves assigning scores to solutions and ordering them accordingly. The diagram also indicates that 30 individuals are rejected and new individuals are added to the population $P_{t+1}$.](image-url)
DTLZ1-5D: $U(x) = \max\{w_1 \times f_1(x), ..., w_5 \times f_5(x)\} \rightarrow \min$

<table>
<thead>
<tr>
<th></th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$w_3$</th>
<th>$w_4$</th>
<th>$w_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTLZ2-5D Cheb. (extreme 1)</td>
<td>0.1</td>
<td>0.15</td>
<td>0.2</td>
<td>0.25</td>
<td>0.3</td>
</tr>
</tbody>
</table>
250 items: $U(x) = \min \{ w_1 x_1(x) - \alpha, w_2 x_2(x) - \alpha \} \rightarrow \max$

Multi-objective Knapsack Problem with 250 items
**XIMEA-DRSA: the explainability issue**

- Consider the 2D knapsack problem with 100 items, $w_2^1 = (1,1)$, and $\alpha=3200$
- The DM is asked every 25 iterations to classify 6 current solutions into *good* or *bad* class
- From this preference information, decision rules are induced to explain the judgments of the DM
- To show how XIMEA-DRSA explains the DM judgments, let’s consider iterations no.: 1, 101, and 576
- Reference solutions:

<table>
<thead>
<tr>
<th>Iteration 1</th>
<th>Iteration 101</th>
<th>Iteration 576</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sol $x^1$</td>
<td>$f_1(\cdot)$</td>
<td>$f_2(\cdot)$</td>
</tr>
<tr>
<td>2908</td>
<td>3002</td>
<td>good</td>
</tr>
<tr>
<td>3048</td>
<td>2906</td>
<td>good</td>
</tr>
<tr>
<td>2890</td>
<td>2991</td>
<td>good</td>
</tr>
<tr>
<td>3042</td>
<td>2868</td>
<td>bad</td>
</tr>
<tr>
<td>2947</td>
<td>2803</td>
<td>bad</td>
</tr>
<tr>
<td>3012</td>
<td>2769</td>
<td>bad</td>
</tr>
</tbody>
</table>

S. Corrente, S. Greco, B. Matarazzo, R. Słowiński: Explainable Interactive Evolutionary Multiobjective Optimization, *OMEGA*, 122 (2024) 102925
**XIMEA-DRSA: the explainability issue**

- **DRSA decision rules induced after the 1\textsuperscript{st} iteration, shown to DM:**
  - **rule 1\_1:** If $f_1(x) \geq 3048$, then $x$ is \textit{good} (supported by $x_2^1$),
  - **rule 2\_1:** If $f_2(x) \geq 2906$, then $x$ is \textit{good} (supported by $x_1^1$, $x_2^1$, and $x_3^1$),
  - **rule 3\_1:** If $f_2(x) \leq 2868$, then $x$ is \textit{bad} (supported by $x_4^1$, $x_5^1$, and $x_6^1$).

- Decision rules are using reduced number of objectives and are \textit{not anonymous}

- While being **transparent** and **intelligible**, the rules used in optimization are also **traceable**

- Reflecting on the decision rules, the **DM learns her preferences**

- **Kahneman’s fast and slow thinking:** rules support slow learning of preferences expressed intuitively by fast decisions
XIMEA-DRSA: the explainability issue

- **Good** decision rules induced after the 101\textsuperscript{st} and 576\textsuperscript{th} iteration:
  - rule 1\_101: If $f_1(x) \geq 3816$ and $f_2(x) \geq 3809$, then $x$ is *good* (supported by $x_{101}^1$)
  - rule 2\_101: If $f_1(x) \geq 3790$ and $f_2(x) \geq 3837$, then $x$ is *good* (supported by $x_{101}^5$)
  - rule 3\_101: If $f_1(x) \geq 3829$ and $f_2(x) \geq 3799$, then $x$ is *good* (supported by $x_{101}^6$)
  - rule 1\_576: If $f_1(x) \geq 3828$ and $f_2(x) \geq 3827$, then $x$ is *good* (supported by $x_{576}^1$)

S. Corrente, S. Greco, B. Matarazzo, R. Słowiński: Explainable Interactive Evolutionary Multiobjective Optimization, *OMEGA*, 122 (2024) 102925
**XIMEA-DRSA: the explainability issue**

- **Decision rules** used to assign a score to the considered solutions induced from classification decisions provided **up to iteration 600**:
  - **rule 1.600**: If $f_1(x) \geq 3799$ and $f_2(x) \geq 3828$, then $x$ is *good* ($\text{born(rule 1.600)} = 126$)
  - **rule 2.600**: If $f_1(x) \geq 3841$ and $f_2(x) \geq 3813$, then $x$ is *good* ($\text{born(rule 2.600)} = 151$)
  - **rule 3.600**: If $f_1(x) \geq 3805$ and $f_2(x) \geq 3820$, then $x$ is *good* ($\text{born(rule 3.600)} = 151$)
  - **rule 4.600**: If $f_1(x) \geq 3831$ and $f_2(x) \geq 3818$, then $x$ is *good* ($\text{born(rule 4.600)} = 151$)
  - **rule 5.600**: If $f_1(x) \geq 3808$ and $f_2(x) \geq 3819$, then $x$ is *good* ($\text{born(rule 5.600)} = 176$)
  - **rule 6.600**: If $f_1(x) \geq 3828$ and $f_2(x) \geq 3827$, then $x$ is *good* ($\text{born(rule 6.600)} = 576$)
  - **rule 7.600**: If $f_2(x) \leq 3727$, then $x$ is *bad* ($\text{born(rule 7.600)} = 51$),
  - **rule 8.600**: If $f_2(x) \leq 3759$, then $x$ is *bad* ($\text{born(rule 8.600)} = 76$),
  - **rule 9.600**: If $f_1(x) \leq 3790$, then $x$ is *bad* ($\text{born(rule 9.600)} = 126$),
  - **rule 10.600**: If $f_2(x) \leq 3785$, then $x$ is *bad* ($\text{born(rule 10.600)} = 126$)
  - **rule 11.600**: If $f_1(x) \leq 3800$, then $x$ is *bad* ($\text{born(rule 11.600)} = 151$)
  - **rule 12.600**: If $f_2(x) \leq 3803$, then $x$ is *bad* ($\text{born(rule 12.600)} = 151$)
  - **rule 13.600**: If $f_1(x) \leq 3805$, then $x$ is *bad* ($\text{born(rule 13.600)} = 176$)
  - **rule 14.600**: If $f_2(x) \leq 3794$, then $x$ is *bad* ($\text{born(rule 14.600)} = 176$)
  - **rule 15.600**: If $f_1(x) \leq 3818$, then $x$ is *bad* ($\text{born(rule 15.600)} = 576$)
  - **rule 16.600**: If $f_2(x) \leq 3819$, then $x$ is *bad* ($\text{born(rule 16.600)} = 576$)

$x=[3828, 3827]$

matches red rules & gets max score

It is also **the best** w.r.t. the true user’s value funct.
Summary and conclusions
Summary and conclusions

- Robust Ordinal Regression is a constructive way of learning DM’s preferences.
- It underlines the evolution of OR and DA towards the AI paradigm of learning.
- It is also a representative of the European School of Decision Aiding, because it goes along with the recommendation of its founder:

  Bernard Roy (1934-2017): „MCDA must be based on models that are co-constructed through interaction with the decision maker. The co-constructed model must be a tool for looking deeper into the subject, exploring, interpreting, debating and even arguing.” (2010)
Thank you for your attention

**Acknowledgment to co-operators:**

Salvatore Corrente, Salvatore Greco, Miłosz Kadziński, Constantin Zopounidis, Yannis Siskos, José Figueira, Jürgen Branke, Vincent Mousseau, Piotr Zielniewicz, Michael Doumpos, …