Knowledge discovery from vague data using dominance-based rough set approach

Roman Słowiński

Institute of Computing Science
Poznań University of Technology

1st Int. Workshop MS2T, Compiègne, Sept. 5, 2013
Plan

- Knowledge discovery
- Main categories of decision problems
- Imperfection of data: rough set concept
- Dominance-based Rough Set Approach (DRSA)
- DRSA for ordinal classification
- DRSA can handle non-ordinal data
- *DRSA for decision under risk & uncertainty
- *DRSA for interactive multiobjective optimization
- Examples of practical applications
- Conclusions & software

* If time permits
Knowledge discovery from data

- The gap between data generation and data comprehension grows up
- *Knowledge Discovery* techniques try to bridge this gap
- Knowledge discovery is an inductive process aiming at identification of:
  - true,
  - non-trivial,
  - useful,
  - directly comprehensible

patterns in data

- Pattern = rule, trend, phenomenon, regularity, anomaly, hypothesis, function etc.

- The patterns are useful for explanation of situations described by data, for prediction of future situations and for building a strategy of intervention
(Ordinal) classification problem (sorting)

DATA

Classes may be ordered:  Class 1 > Class 2 > ... > Class p
Ranking problem

DATA

Partial or complete ranking of objects
Choice problem (multiobjective optimization)

DATA

Chosen subset of best objects

Rejected subset of objects
Knowledge discovery from data

- Data about reality is **imperfect** for various reasons:
  - inexact determination of some features,
  - uncertainty of some features due to their unpredictable behavior,
  - imprecise measurement of some features,
  - presence of humans with their subjective and unstable judgments,
  - finite language of description = granularity of information,
  - ambiguity and inconsistency of information
  - ...

- All this makes that our reasoning about reality through various models is only approximate and bounded by conclusions which are certainly true or certainly false according to the available data.
Adopted perspective

- **Rough Set** as a concept useful for *structuring granular and inconsistent information* prior to *induction of patterns*

- We will concentrate on rough set reasoning about data describing classification data

- Moreover, as decisions involve evaluation (*order, quality*), we admit that *value sets of attributes may be ordered*
Relation to known paradigms

- AI and Machine Learning: “Learning from examples”
- Decision aiding & Recommender systems: “Preference learning”
- Interacting systems: “Man-machine interaction”, “Multi-agent systems”
- Operations Research: “Analytics - the scientific process of transforming data into insight for making better decisions”
Classification data – example of technical diagnostics

- 176 buses (objects)
- 8 symptoms (attributes)
- Decision = technical state:
  - 3 – good state (in use)
  - 2 – minor repair
  - 1 – major repair (out of use)
- Knowledge discovery = finding relationships between symptoms & technical state
- Knowledge explains expert’s decisions and supports diagnosis of new buses

<table>
<thead>
<tr>
<th>Attributes: 9 of 10</th>
<th>Examples: 76</th>
<th>Decision: State</th>
<th>Missing Values: No</th>
</tr>
</thead>
<tbody>
<tr>
<td>MaxSpeed</td>
<td>ComprPressure</td>
<td>Blacking</td>
<td>Torque</td>
</tr>
<tr>
<td>1. 90</td>
<td>2</td>
<td>38</td>
<td>481</td>
</tr>
<tr>
<td>2. 76</td>
<td>2</td>
<td>70</td>
<td>420</td>
</tr>
<tr>
<td>......</td>
<td>......</td>
<td>......</td>
<td>......</td>
</tr>
<tr>
<td>27. 68</td>
<td>2</td>
<td>70</td>
<td>400</td>
</tr>
</tbody>
</table>
Representation of knowledge

- **Scoring function:**
  \[ U(a) = \sum_{i=1}^{n} k_i g_i(a) \] or
  \[ U(a) = \sum_{i=1}^{n} u_i [g_i(a)] \]
  
  like in MAUT, Discriminant Analysis, Logistic Regression or Perceptron,

  e.g.
  \[ U(a) = 0.21 \times g_{\text{Speed}}(a) + 0.03 \times g_{\text{Compr}}(a) + ... + 0.18 \times g_{\text{Power}}(a) = 0.45 \]
Representation of knowledge

- Scoring function: \( U(a) = \sum_{i=1}^{n} k_i g_i(a) \) or \( U(a) = \sum_{i=1}^{n} u_i [g_i(a)] \)

like in MAUT, Discriminant Analysis, Logistic Regression or Perceptron,

\[ U(a) = 0.21 \times g_{\text{Speed}}(a) + 0.03 \times g_{\text{Compr}}(a) + \ldots + 0.18 \times g_{\text{Power}}(a) = 0.45 \]

\[ \sum_{i=1}^{n} a_i g_k(a) = U(a) \]

State 1: 0.0
State 2: 0.34
State 3: 0.76
State 4: 1.0

\( U(a) \)
Representation of knowledge

- **Scoring function:**  \( U(a) = \sum_{i=1}^{n} k_i g_i(a) \)  or  \( U(a) = \sum_{i=1}^{n} u_i [g_i(a)] \)

  like in MAUT, Discriminant Analysis, Logistic Regression or Perceptron,

  e.g.  \( U(a) = 0.21 \times g_{\text{Speed}}(a) + 0.03 \times g_{\text{Compr}}(a) + \ldots + 0.18 \times g_{\text{Power}}(a) = 0.45 \)

- **Decision rules or trees,**

  like in Artificial Intelligence, Data Mining or Learning from Examples,

  e.g.  \( \text{if} \ \text{OilCons} \leq 1 \ \& \ \text{WinterGasCons} \leq 25, \ \text{then} \ \text{State} \geq 2 \)

  \( \text{if} \ \text{MaxSpeed} \leq 85 \ \& \ \text{WinterGasCons} \geq 25, \ \text{then} \ \text{State} \leq 2 \)
Representation of knowledge

- **Scoring function:** \( U(a) = \sum_{i=1}^{n} k_i g_i(a) \) or \( U(a) = \sum_{i=1}^{n} u_i [g_i(a)] \)

  like in MAUT, Discriminant Analysis, Logistic Regression or Perceptron,
  e.g. \( U(a) = 0.21 \times g_{\text{Speed}}(a) + 0.03 \times g_{\text{Compr}}(a) + \ldots + 0.18 \times g_{\text{Power}}(a) = 0.45 \)

- **Decision rules or trees,**
  like in Artificial Intelligence, Data Minining or Learning from Examples,
  e.g. if OilCons \( \leq 1 \) & WinterGasCons \( \leq 25 \), then State \( \geq 2 \)
  if MaxSpeed \( \leq 85 \) & WinterGasCons \( \geq 25 \), then State \( \leq 2 \)

- „People make decisions by searching for rules that provide good justification of their choices“ (Slovic, 1975)
Aggregation of multiple attribute (criteria) evaluations

- Three families of **aggregation methods (preference modelling)**:
  - **Multiple Attribute Utility Theory (MAUT)** using a value function,
    
    \[ U(a) = \sum_{i=1}^{n} k_i g_i(a), \quad U(a) = \sum_{i=1}^{n} u_i[g_i(a)], \]
    
    Choquet/Sugeno integral
  
  - **Outranking methods** using an outranking relation \( S \)
    
    \[ a S b = \text{“}a \text{ is at least as good as } b\text{”} \]
  
  - **Decision rule approach** using a set of „if..., then...“ decision rules

- Decision rule model is the most general of all three

Syntax of decision rules

**Ordinal classification**

- If \( x_{q1} \preceq_{q1} r_{q1} \) and \( x_{q2} \preceq_{q2} r_{q2} \) and ... \( x_{qp} \preceq_{qp} r_{qp} \), then \( x \rightarrow \text{class } t \) or better
- If \( x_{q1} \succeq_{q1} r_{q1} \) and \( x_{q2} \succeq_{q2} r_{q2} \) and ... \( x_{qp} \succeq_{qp} r_{qp} \), then \( x \rightarrow \text{class } t \) or worse

**Choice ranking**

- If \( (x \succ_{q1} \succeq_{q1} h(q1)) \) and \( (x \succ_{q2} \succeq_{q2} h(q2)) \) and ... \( (x \succ_{qp} \succeq_{qp} h(qp)) \), then \( x \text{Sy} \)
- If \( (x \preceq_{q1} \succeq_{q1} h(q1)) \) and \( (x \preceq_{q2} \succeq_{q2} h(q2)) \) and ... \( (x \preceq_{qp} \succeq_{qp} h(qp)) \), then \( x \text{Scy} \)

**Cardinal criteria**

- If \( x_{g1} \succeq_{g1} r_{q1} \) and \( y_{g1} \preceq_{g1} r'_{q1} \) and ... \( x_{gp} \succeq_{gp} r_{gp} \) and \( y_{gp} \preceq_{gp} r'_{gp} \), then \( x \text{Sy} \)
- If \( x_{g1} \preceq_{g1} r_{q1} \) and \( y_{g1} \succeq_{g1} r'_{q1} \) and ... \( x_{gp} \preceq_{gp} r_{gp} \) and \( y_{gp} \succeq_{gp} r'_{gp} \), then \( x \text{Scy} \)

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

Why should we seek for rules rather than for a real-valued function?

- Description of complex phenomena by recursive estimation techniques applied on historical data (Int. J. Environment and Pollution, vol.12, no.2/3, 1999)

- The dependence of the size of the mouth of a river in month $k$, represented by the relative tidal energy ($RTE_k$), from $RTE_{k-1}$, the river flow ($F_{k-1}$), the onshore wind ($W_{k-1}$) and the crude monthly count of storm events ($S_k$) (Elford et al. 1999; Murray Mouth, Australia):

$$RTE_k = A_1RTE_{k-1} + A_2 \frac{(F_{k-1} - 200)^{2.4}}{8RTE_{k-1} + 1} + A_3 \frac{W_{k-1}}{8RTE_{k-1} + 1} + A_4S_k + \varepsilon_k$$

where the exponent 2.4 was tuned by „trial and error“, coefficients $A_1, A_2, A_3, A_4$ were computed using a recursive least squares (RLS), and $\varepsilon_k$ is the model error.
Why should we seek for rules rather than for a real-valued function?

- Description of complex phenomena by recursive estimation techniques applied on historical data (Int. J. Environment and Pollution, vol.12, no.2/3, 1999)

- The **impact of urban stormwater on the quality of the receiving water** (Rossi, Słowiński, Susmaga 1999; Lausanne and Genève).

- Example of rule induced from empirical observation of some sites:

  If the site is of type 2 (residential), and total rainfall is up to 8 mm, and max intensity of rain is between 2.7 and 11.2 mm/h, then total mass of suspended solids is < 2.2 kg/ha

- The rule is more **expressive** and involves **heterogeneous data**: nominal, qualitative and quantitative.
Rough set concept
Inconsistencies in data – Rough Set Theory

- Zdzisław Pawlak (1926–2006)

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics</th>
<th>Physics</th>
<th>Literature</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>
Inconsistencies in data – Rough Set Theory

- The granules of indiscernible objects are used to approximate classes

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics (M)</th>
<th>Physics (Ph)</th>
<th>Literature (L)</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>
Inconsistencies in data – Rough Set Theory

- Lower approximation of class “good”

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics (M)</th>
<th>Physics (Ph)</th>
<th>Literature (L)</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>
Inconsistencies in data – Rough Set Theory

- Lower and upper approximation of class „good”

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics (M)</th>
<th>Physics (Ph)</th>
<th>Literature (L)</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>
Relationship between RST and DST

- Let $\mathbf{Cl}$ be the frame of discernment compatible with the decision table $S = \langle U, C \cup \{d\} \rangle$, let also $T(\theta) = \{t : Cl_t \in \mathbf{Cl}\}$, where $\theta \subseteq \mathbf{Cl}$

- For any $\theta \subseteq \mathbf{Cl}$ the belief function can be calculated as:

$$Bel_S(\theta) = \sum_{\Delta \in \theta} m_S(\Delta) = \sum_{\Delta \in \theta} \frac{\text{card}(\{x \in U : \delta_S(x) = \Delta\})}{\text{card}(U)}$$

$$= \sum_{t \in T(\theta)} \frac{\text{card}(C(Cl_t))}{\text{card}(U)} + \sum_{\substack{\Delta \in \theta \\ \text{card}(\Delta) > 1}} \frac{\text{card}(\{x \in U : \delta_S(x) = \Delta\})}{\text{card}(U)} = \frac{\text{card}(C(\bigcup_{t \in T(\theta)} Cl_t))}{\text{card}(U)}$$

where $\delta_S(x) = \{Cl_t : \exists y \in U, y \in I_C(x) \text{ and } y \in Cl_t, t \in T(\mathbf{Cl})\}$

is called generalized decision for object $x$ (cluster of classes, with no possibility of discernment using knowledge about $S = \langle U, C \cup \{d\} \rangle$)
For any $\theta \subseteq CI$ the plausibility function can be calculated as:

$$P_{I_S}(\theta) = 1 - Bel_{I_S}(CI - \theta) = 1 - \frac{\text{card}(\bigcup_{t \in T(\text{CI-}\theta)} Cl_t)}{\text{card}(U)} = \frac{\text{card}(\bigcup_{t \in T(\theta)} Cl_t)}{\text{card}(U)}$$
Consistency measures – probabilistic rough sets

- Gain-type consistency measure:
  - Rough membership, \( \mu \)-consistency measure

  \[
  \mu_{Cl_t}^p (x) = \frac{|D_p^+ (x) \cap Cl_t^{\geq}|}{|D_p^+ (x)|} \quad \mu_{Cl_t}^p (x) = \frac{|D_p^- (x) \cap Cl_t^{\leq}|}{|D_p^- (x)|}
  \]

- It can be interpreted as an estimate of conditional probability:

  \[
  Pr\left(y \in Cl_t^{\geq} \mid y \in D_p^+ (x)\right) \quad Pr\left(y \in Cl_t^{\leq} \mid y \in D_p^- (x)\right)
  \]


Cost-type consistency measure:

\( \varepsilon \)-consistency measure

\[
\varepsilon^P_{Cl^\geq_t} (x) = \frac{|D^+_p(x) \cap -Cl^\geq_t|}{|\neg Cl^\geq_t|}
\]

\[
\varepsilon^P_{Cl^\leq_t} (x) = \frac{|D^-_p(x) \cap -Cl^\leq_t|}{|\neg Cl^\leq_t|}
\]

It can be interpreted as an estimate of conditional probability:

\[
Pr\left(y \in D^+_p(x) \mid y \in \neg Cl^\geq_t \right)
\]

\[
Pr\left(y \in D^-_p(x) \mid y \in \neg Cl^\leq_t \right)
\]

The intuition behind \( \varepsilon \)-consistency measure: it says how far the implications \( y \in D^+_p(x) \Rightarrow y \in Cl^\geq_t, \ y \in D^-_p(x) \Rightarrow y \in Cl^\leq_t \) are not supported by the data.
**IRSA – rules induced from rough approximations**

- **Certain decision rule** supported by objects from the **lower approximation** of class „good” (discriminant rule)

  \[
  \text{If } \text{Lit} = \text{good}, \text{ then } \text{Student is certainly good} \\
  \{S5, S6\}
  \]

- **Possible decision rule** supported by objects from the **upper approximation** of class „good” (partly discriminant rule)

  \[
  \text{If } \text{Phys} = \text{good}, \text{ then } \text{Student is possibly good} \\
  \{S3, S4, S6\}
  \]

- **Approximate decision rule** supported by objects from the **boundary** of class „medium” or „good”

  \[
  \text{If } \text{Phys} = \text{good} \& \text{Lit} = \text{medium}, \text{ then } \text{Student is medium or good} \\
  \{S3, S4\}
  \]
What is missing to Indiscernibility-based Rough Set Approach?

- Classical rough set approach does not detect inconsistency w.r.t. dominance (Pareto principle) – a basic principle in decision making

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics (M)</th>
<th>Physics (Ph)</th>
<th>Literature (L)</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>medium</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>
Certain decision rules based on indiscernibility are inconsistent with respect to the dominance principle (monotonicity constraints):

If Math=good & Lit=bad, then Student is certainly bad \{S1\}

If Math=medium & Lit=bad, then Student is certainly medium \{S2\}
DRSA – example of technical diagnostics

- 176 vehicles (objects)
- 8 symptoms
- decision = technical state:
  - 3 – good state (in use)
  - 2 – minor repair
  - 1 – major repair (out of use)
- there is a monotonic relationship between each symptom and the decision
- inconsistent objects: 11, 12, 39

<table>
<thead>
<tr>
<th></th>
<th>MaxSpeed</th>
<th>CompPressure</th>
<th>Blacking</th>
<th>Torque</th>
<th>SummerCons</th>
<th>WinterCons</th>
<th>OilCons</th>
<th>HorsePower</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>90</td>
<td>2</td>
<td>38</td>
<td>481</td>
<td>21</td>
<td>26</td>
<td>0</td>
<td>145</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>76</td>
<td>2</td>
<td>70</td>
<td>420</td>
<td>22</td>
<td>25</td>
<td>2</td>
<td>110</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>63</td>
<td>1</td>
<td>82</td>
<td>400</td>
<td>22</td>
<td>24</td>
<td>3</td>
<td>101</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>90</td>
<td>2</td>
<td>49</td>
<td>477</td>
<td>21</td>
<td>25</td>
<td>1</td>
<td>138</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>85</td>
<td>2</td>
<td>52</td>
<td>460</td>
<td>21</td>
<td>25</td>
<td>1</td>
<td>130</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>72</td>
<td>2</td>
<td>73</td>
<td>425</td>
<td>23</td>
<td>27</td>
<td>2</td>
<td>112</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>98</td>
<td>2</td>
<td>50</td>
<td>490</td>
<td>21</td>
<td>24</td>
<td>1</td>
<td>140</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>87</td>
<td>2</td>
<td>56</td>
<td>465</td>
<td>22</td>
<td>27</td>
<td>1</td>
<td>135</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>90</td>
<td>2</td>
<td>16</td>
<td>486</td>
<td>26</td>
<td>27</td>
<td>0</td>
<td>150</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>60</td>
<td>1</td>
<td>95</td>
<td>400</td>
<td>23</td>
<td>24</td>
<td>4</td>
<td>96</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>80</td>
<td>2</td>
<td>60</td>
<td>451</td>
<td>21</td>
<td>26</td>
<td>1</td>
<td>125</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>78</td>
<td>2</td>
<td>63</td>
<td>448</td>
<td>21</td>
<td>26</td>
<td>1</td>
<td>120</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>90</td>
<td>2</td>
<td>26</td>
<td>482</td>
<td>22</td>
<td>24</td>
<td>0</td>
<td>148</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>62</td>
<td>1</td>
<td>93</td>
<td>400</td>
<td>22</td>
<td>28</td>
<td>3</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>82</td>
<td>2</td>
<td>54</td>
<td>461</td>
<td>22</td>
<td>26</td>
<td>1</td>
<td>132</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>65</td>
<td>2</td>
<td>67</td>
<td>402</td>
<td>22</td>
<td>23</td>
<td>2</td>
<td>103</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>90</td>
<td>2</td>
<td>51</td>
<td>468</td>
<td>22</td>
<td>26</td>
<td>1</td>
<td>138</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>90</td>
<td>2</td>
<td>15</td>
<td>488</td>
<td>20</td>
<td>23</td>
<td>0</td>
<td>150</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td>76</td>
<td>2</td>
<td>65</td>
<td>428</td>
<td>27</td>
<td>33</td>
<td>2</td>
<td>116</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>85</td>
<td>2</td>
<td>50</td>
<td>454</td>
<td>21</td>
<td>26</td>
<td>1</td>
<td>129</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>85</td>
<td>2</td>
<td>58</td>
<td>450</td>
<td>22</td>
<td>25</td>
<td>1</td>
<td>126</td>
<td>2</td>
</tr>
<tr>
<td>22</td>
<td>88</td>
<td>2</td>
<td>48</td>
<td>458</td>
<td>22</td>
<td>25</td>
<td>1</td>
<td>130</td>
<td>3</td>
</tr>
<tr>
<td>23</td>
<td>60</td>
<td>1</td>
<td>90</td>
<td>400</td>
<td>24</td>
<td>28</td>
<td>4</td>
<td>95</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>64</td>
<td>2</td>
<td>71</td>
<td>420</td>
<td>23</td>
<td>25</td>
<td>2</td>
<td>105</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>75</td>
<td>2</td>
<td>64</td>
<td>432</td>
<td>22</td>
<td>25</td>
<td>1</td>
<td>114</td>
<td>2</td>
</tr>
<tr>
<td>26</td>
<td>74</td>
<td>2</td>
<td>64</td>
<td>420</td>
<td>21</td>
<td>25</td>
<td>1</td>
<td>110</td>
<td>2</td>
</tr>
<tr>
<td>27</td>
<td>68</td>
<td>2</td>
<td>70</td>
<td>400</td>
<td>22</td>
<td>26</td>
<td>2</td>
<td>100</td>
<td>1</td>
</tr>
</tbody>
</table>

Attributes: 9 of 10  Examples: 76  Decision: State  Missing Values: No
Dominance-based Rough Set Approach: DRSA
Classical Rough Set Theory vs. Dominance-based Rough Set Theory

**Classical Rough Set Theory**

\[ \Downarrow \]

**Indiscernibility principle**

If \( x \) and \( y \) are indiscernible with respect to all relevant attributes, then \( x \) should classified to the same class as \( y \)

**Dominance-based Rough Set Theory**

\[ \Downarrow \]

**Dominance principle** (monotonicity constraints)

If \( x \) is at least as good as \( y \) with respect to all relevant criteria, then \( x \) should be classified at least as good as \( y \)

Dominance principle as monotonicity constraint principle

- **Dominance-based Rough Set Approach (DRSA)** permits representation and analysis of **all phenomena involving monotonicity relationship** between specific **measures** or **perceptions**, e.g.

  "the more a tomato is red, and the more it is soft, the more it is ripe"
  "the older the car, the more likely its breakdown" *

  or

  "the more similar are the causes, the more similar are the effects one can expect"**


Monotonicity and induction

- „The procedure of induction consists in accepting as true the simplest law that can be reconciled with our experiences”
  (L. Wittgenstein, Tractatus Logico-Philosophicus, 6.363)
- This simplest law is just monotonicity and, therefore, inductive discovery of rules can be seen as a specific way of dealing with monotonicity
- Dominance-based Rough Set concept permits data structuring wrt possible violation of dominance (lower appx, upper appx, boundary) prior to rule induction

Dominance-based Rough Set Approach (DRSA)

- In order to handle **monotonic dependency** between conditions and decision (class assignment):

  \[ Cl_t^\geq = \bigcup_{s \geq t} Cl_s \]  
  \[ \text{upward union of classes, } t=2,\ldots,m \ ("at least" \text{ class } Cl_t) \]

  \[ Cl_t^\leq = \bigcup_{s \leq t} Cl_s \]  
  \[ \text{downward union of classes, } t=1,\ldots,m-1 \ ("at most" \text{ class } Cl_t) \]

- \( Cl_t^\geq \) and \( Cl_t^\leq \) are positive and negative **dominance cones** in decision space reduced to single dimension

Ordered classes

\[ Cl_6 \]
\[ Cl_5 \]
\[ Cl_4 \]
\[ Cl_3 \]
\[ Cl_2 \]
\[ Cl_1 \]

\[ Cl_4^\geq = \{ Cl_4, Cl_5, Cl_6 \} \]

\[ Cl_3^\leq = \{ Cl_3, Cl_2, Cl_1 \} \]
Dominance-based Rough Set Approach (DRSA)

- $D_P$ – dominance relation (partial preorder) in condition space, $P \subseteq C$
- Granules of knowledge are dominance cones in condition space
  \[
  D_P^+(x) = \{ y \in U : y D_P x \} : \text{P-dominating set (positive cone)}
  \]
  \[
  D_P^-(x) = \{ y \in U : x D_P y \} : \text{P-dominated set (negative cone)}
  \]
- Classification patterns to be discovered are functions representing granules $Cl_t^\geq, Cl_t^\leq$, by granules $D_P^+(x), D_P^-(x)$
Dominance cones wrt object $x$ – example ($C_1 \prec C_2 \prec C_3$)
Dominance cones wrt object $x$ – example ( $Cl_1 \prec Cl_2 \prec Cl_3$ )
Lower approximations of “at most $Cl_1$” and “at least $Cl_2$”
Lower approximations of „at most $Cl_2$“ and „at least $Cl_3$“

\[ Cl_1 \prec Cl_2 \prec Cl_3 \]
Dominance-based Rough Set Approach vs. Classical RSA

Comparison of CRSA and DRSA

Classes: ▲ > ● ≥ □

\[ P(X) = \{ x \in U : I_p(x) \subseteq X \} \]

\[ \overline{P}(X) = \bigcup_{x \in X} I_p(x) \]

\[ P(Cl_t^{\ge}) = \{ x \in U : D_p^+ (x) \subseteq Cl_t^{\ge} \} \]

\[ \overline{P}(Cl_t^{\ge}) = \bigcup_{x \in Cl_t^{\ge}} D_p^+(x) \]
DRSA for **multiple criteria classification**

- Example of preference information about students:

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics (M)</th>
<th>Physics (Ph)</th>
<th>Literature (L)</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>good</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>

- Examples of classification of S1 and S2 are inconsistent

Quality of approximation by \( \{M, Ph, L\} = 6/8 = 0.75 \)
DRSA for **multiple criteria classification**

- If we eliminate **Literature**, then more inconsistencies appear:

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics (M)</th>
<th>Physics (Ph)</th>
<th>Literature (L)</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>good</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>

- Examples of classification of **S1, S2, S3 and S5** are inconsistent
DRSA for multiple criteria classification

- Elimination of Mathematics does not increase inconsistencies:

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics (M)</th>
<th>Physics (Ph)</th>
<th>Literature (L)</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>good</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>

- Subset of criteria \{Ph,L\} is a reduct of \{M,Ph,L\}
DRSA for multiple criteria classification

- Elimination of Physics also does not increase inconsistencies:

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics (M)</th>
<th>Physics (Ph)</th>
<th>Literature (L)</th>
<th>Overall class</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>good</td>
<td>medium</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S2</td>
<td>medium</td>
<td>medium</td>
<td>bad</td>
<td>medium</td>
</tr>
<tr>
<td>S3</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
<td>medium</td>
</tr>
<tr>
<td>S4</td>
<td>good</td>
<td>good</td>
<td>medium</td>
<td>good</td>
</tr>
<tr>
<td>S5</td>
<td>good</td>
<td>medium</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S6</td>
<td>good</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>S7</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
<td>bad</td>
</tr>
<tr>
<td>S8</td>
<td>bad</td>
<td>bad</td>
<td>medium</td>
<td>bad</td>
</tr>
</tbody>
</table>

- Subset of criteria \{M,L\} is a reduct of \{M,Ph,L\}
- Intersection of reducts \{M,L\} and \{Ph,L\} gives the core \{L\}
DRSA for multicriteria classification

- Quality of approximation: $\gamma_p(CI) = 0.75$
- Reducts: $\{M,L\}$ and $\{Ph,L\}$
- Core: $\{L\}$

Set of decision rules in terms of $\{M,Ph,L\}$ representing preferences:

*If* $M \succ$ good & $L \succ$ medium, then student $\succ$ good \{S4,S5,S6\}

*If* $M \succ$ medium & $L \succ$ medium, then student $\succ$ medium \{S3,S4,S5,S6\}

*If* $M \succ$ medium & $L \preceq$ bad, then student is bad or medium \{S1,S2\}

*If* $Ph \preceq$ medium & $L \preceq$ medium then student $\preceq$ medium \{S1,S2,S3,S7,S8\}

*If* $M \preceq$ bad, then student $\preceq$ bad \{S7,S8\}

Fuzzy measure with Choquet capacity

meaningful for the role of criteria
Partial profiles of rules are matching object profiles if dominance holds.

Rules $r_2$ and $r_3$ are matching object $a$. 
Using DRSA rules as a decision model

- New student to be evaluated

<table>
<thead>
<tr>
<th>Student</th>
<th>Mathematics</th>
<th>Physics</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>S9</td>
<td>medium</td>
<td>medium</td>
<td>good</td>
</tr>
</tbody>
</table>

- Set of activated decision rules:

  \( \text{If } M \geq \text{medium } \& L \geq \text{medium, then } \text{student } \geq \text{medium} \) \{S3,S4,S5,S6\}

  \( \text{If } M \leq \text{medium, then } \text{student } \leq \text{medium} \) \{S2,S3,S7,S8\}

- Set of non-activated decision rules:

  \( \text{If } M \geq \text{good } \& L \geq \text{medium, then } \text{student } \geq \text{good} \) \{S4,S5,S6\}

  \( \text{If } M \geq \text{medium } \& L \leq \text{bad, then } \text{student is bad or medium} \) \{S1,S2\}

  \( \text{If } M \leq \text{bad, then } \text{student } \leq \text{bad} \) \{S7,S8\}
Important feature of DRSA

- DRSA exploits ordinal information only, and decision rules do not convert ordinal information into numeric one.

- “Si l’ordre apparaît quelque part dans la qualité, pourquoi chercherions-nous à passer par l’intermédiaire du nombre?” (G. Bachelard 1934)

  (“If an order appears somewhere in quality, why should we like to interpret this order through numerical values?”)

- Pareto-dominance can be replaced by Lorenz-dominance, making decision rules more equitable and risk-averse.
Preference modelling by dominance-based decision rules

- Dominance-based „if..., then...” decision rules are the only aggregation operators that:
  - give account of most complex interactions among criteria,
  - 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 criteria and a comprehensive judgment

DRSA can also handle non-ordinal data
DRSA for classification with unknown monotonicity constraints

- Attributes with unknown monotonic relationship w.r.t. decision

1. **Ordinal** (number-coded) attributes
   - qualitative (small (1), medium (2), ..., large (k): e.g., size)
   - quantitative (numerical: e.g., temperature)

   Each ordinal attribute $a_i$ is replaced by 2 criteria:
   - gain-type criterion $q'_i$ and cost-type criterion $q''_i$

2. **Nominal** (not ordered) attributes (blue, red, ..., white: e.g., color)

   Each nominal attribute $a_i$ (taking 1 of $k$ values, $k > 2$)
   is replaced by $2 \times k$ binary criteria: for each $h \in \{1, \ldots, k\}$,
   - gain-type 0-1 criterion $q'_i(h)$ and cost-type 0-1 criterion $q''_i(h)$
DRSA for classification with unknown monotonicity constraints

- Decision attribute $d$ makes partition of $U$ into a finite number of non-ordered decision classes $\mathcal{Cl} = \{ Cl_t, \ t=1,...,m \}$

- Using DRSA, one approximates:
  - in case of $m=2$ (binary classification): $Cl_1$ and $\neg Cl_1 = Cl_2$
  - in case of $m>2$: $Cl_t$ and $\neg Cl_t$, for each $t \in \{1,...,k\}$, i.e.
Example of application of DRSA to non-ordinal data

Set of patients after radical prostatectomy.

<table>
<thead>
<tr>
<th>Id</th>
<th>Age</th>
<th>Gleason</th>
<th>PSA</th>
<th>Volume</th>
<th>Recurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>10</td>
<td>2.0</td>
<td>large</td>
<td>other</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>7</td>
<td>1.2</td>
<td>large</td>
<td>local</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>4</td>
<td>0.1</td>
<td>medium</td>
<td>local</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>2</td>
<td>0.8</td>
<td>medium</td>
<td>no</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>3</td>
<td>0.3</td>
<td>small</td>
<td>local</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>3</td>
<td>0.3</td>
<td>small</td>
<td>no</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>7</td>
<td>0.5</td>
<td>small</td>
<td>no</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>5</td>
<td>0.4</td>
<td>small</td>
<td>no</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>2</td>
<td>0.5</td>
<td>small</td>
<td>no</td>
</tr>
<tr>
<td>10</td>
<td>40</td>
<td>4</td>
<td>0.5</td>
<td>small</td>
<td>no</td>
</tr>
</tbody>
</table>
Example of application of DRSA to non-ordinal data

Transformed set of patients after radical prostatectomy—binary classification into “no” and “→ no”.

<table>
<thead>
<tr>
<th>Id</th>
<th>Age'</th>
<th>Age''</th>
<th>Gleason'</th>
<th>Gleason''</th>
<th>PSA'</th>
<th>PSA''</th>
<th>V-s'</th>
<th>V-s''</th>
<th>V-m'</th>
<th>V-m''</th>
<th>V-1'</th>
<th>V-1''</th>
<th>R-no</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>60</td>
<td>10</td>
<td>10</td>
<td>2.0</td>
<td>2.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>20</td>
<td>7</td>
<td>7</td>
<td>1.2</td>
<td>1.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>40</td>
<td>4</td>
<td>4</td>
<td>0.2</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>45</td>
<td>2</td>
<td>2</td>
<td>0.8</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>50</td>
<td>3</td>
<td>3</td>
<td>0.3</td>
<td>0.3</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>6</td>
<td>50</td>
<td>50</td>
<td>3</td>
<td>3</td>
<td>0.3</td>
<td>0.3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>40</td>
<td>7</td>
<td>7</td>
<td>0.6</td>
<td>0.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>25</td>
<td>5</td>
<td>5</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>25</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>40</td>
<td>40</td>
<td>4</td>
<td>4</td>
<td>0.5</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Example of application of DRSA to non-ordinal data

- Two decision rules are sufficient to cover all consistent objects from the table with binary classification „no” and „¬no” for recurrence

  1. \( \text{if } \text{Gleason}'' \geq 4 \text{ and } V-s'' \leq 0, \text{ then } R-no \leq 0, \)

  2. \( \text{if } PSA' \geq 0.4 \text{ and } PSA'' \leq 0.8, \text{ then } R-no \geq 1. \)

- Elementary condition \( V-s' \leq 0 \) from the rule 1) is be read as: „Volume is not small”. After returning to original scales:

  1. \( \text{if } \text{Gleason} \geq 4 \text{ and } \text{Volume} \in \{\text{medium, large}\}, \text{ then } \text{Recurrence is } \neg \text{no}, \)

  2. \( \text{if } PSA \in [0.4,0.8], \text{ then } \text{Recurrence is no.} \)
Example of application of DRSA to non-ordinal data

- Two decision rules are sufficient to cover all consistent objects from the table with binary classification "local" and "¬local" for recurrence

3: if Age ≥ 25 and PSA ≥ 0.4, then Recurrence is ¬local,

4: if Age ≤ 40 and Volume ∈ {medium, large}, then Recurrence is local.

- Other two rules are sufficient to cover all consistent objects from the table with binary classification "other" and "¬other" for recurrence

5: if PSA ≤ 1.2, then Recurrence is ¬other,

6: if PSA ≥ 2, then Recurrence is other.
Example of application of DRSA to non-ordinal data

- **Classification** of patient \((x_{11})\) using the six rules

<table>
<thead>
<tr>
<th>Id</th>
<th>Age'</th>
<th>Age''</th>
<th>Gleason'</th>
<th>Gleason''</th>
<th>PSA'</th>
<th>PSA''</th>
<th>V-s'</th>
<th>V-s''</th>
<th>V-m'</th>
<th>V-m''</th>
<th>V-l'</th>
<th>V-l''</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>30</td>
<td>30</td>
<td>2</td>
<td>2</td>
<td>0.6</td>
<td>0.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- The patient is covered by the following rules:
  - rule 2, suggesting assignment to class “no”,
  - rule 3, dissuading assignment to class “local” (i.e. suggesting assignment to “\(\neg\) local”),
  - rule 5, dissuading assignment to class “other” (i.e. suggesting assignment to “\(\neg\) other”).

- The result of classification is as follows:

  \[
  \text{score}_{r_{\text{no}}} (\text{no}, x_{11}) = \frac{5^2}{5 \times 5} = 1, \\
  \text{score}_{r_{\neg\text{local}}} (\neg\text{local}, x_{11}) = \frac{6^2}{6 \times 6} = 1, \\
  \text{score}_{r_{\neg\text{other}}} (\neg\text{other}, x_{11}) = \frac{9^2}{9 \times 9} = 1, \\
  \]

  - \(\text{score}(\text{no}, x_{11}) = 1\),
  - \(\text{score}(\text{local}, x_{11}) = -1\),
  - \(\text{score}(\text{other}, x_{11}) = -1\).

No recurrence for \(x_{11}\)
DRSA for multicriteria choice and ranking
Rough approximation of binary relations: DRSA for multicriteria choice and ranking

- Preference information of the DM in the form of pairwise comparisons of reference objects is put in a **pairwise comparison table (PCT)**

- Comparing objects \(a, b \in A^R\) on
  - a **cardinal** criterion, one puts in PCT the value \(\Delta_i(a, b) = g_i(a) - g_i(b)\)
  - an **ordinal** criterion, one puts in PCT the ordered pair \((g_i(a), g_i(b))\)

### Pairwise Comparison Table (PCT)

<table>
<thead>
<tr>
<th>Pair of reference objects</th>
<th>Evaluations on criteria</th>
<th>Preference information</th>
</tr>
</thead>
<tbody>
<tr>
<td>((a, b))</td>
<td>(\Delta_i(a, b))</td>
<td>((g_n(a), g_n(b)))</td>
</tr>
<tr>
<td>((b, a))</td>
<td>(\Delta_i(b, a))</td>
<td>((g_n(b), g_n(a)))</td>
</tr>
<tr>
<td>((b, c))</td>
<td>(\Delta_i(b, c))</td>
<td>((g_n(b), g_n(c)))</td>
</tr>
<tr>
<td>((d, e))</td>
<td>(\Delta_i(d, e))</td>
<td>((g_n(d), g_n(e)))</td>
</tr>
</tbody>
</table>

\(B \subseteq A^R \times A^R\)

- **\(S\) – outranking**
- **\(S^c\) – non-outranking**

\(G = \{g_1, \ldots, g_n\}\)

\(g_1\)-cardinal; \(g_n\)-ordinal
**Syntax** of decision rules

**Cardinal criteria**

If \((x \succ^{h(q1)}_{q1} y)\) and \((x \succ^{h(q2)}_{q2} y)\) and ... \((x \succ^{h(qp)}_{qp} y)\), then \(xSy\)

If \((x \preceq^{h(q1)}_{q1} y)\) and \((x \preceq^{h(q2)}_{q2} y)\) and ... \((x \preceq^{h(qp)}_{qp} y)\), then \(xS^c y\)

**Ordinal criteria**

If \(x_{g1} \succeq_{g1} r_{q1} \) and \(y_{g1} \preceq_{g1} r'_{q1}\) and ... \(x_{gp} \succeq_{gp} r_{gp}\) and \(y_{gp} \preceq_{gp} r'_{gp}\), then \(xSy\)

If \(x_{g1} \preceq_{g1} r_{q1} \) and \(y_{g1} \succeq_{g1} r'_{q1}\) and ... \(x_{gp} \preceq_{gp} r_{gp}\) and \(y_{gp} \succeq_{gp} r'_{gp}\), then \(xS^c y\)

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

---

Decision rule approach to decision under risk & uncertainty
DRSA for decision under risk and uncertainty

- $ST=\{st_1, st_2, st_3, \ldots\}$ – set of elementary states of the world

- $Pr$ – a priori probability distribution over $ST$
  e.g.: $pr_1=0.25$, $pr_2=0.30$, $pr_3=0.35$, ...

- $A=\{A_1, A_2, A_3, A_4, A_5, A_6, \ldots\}$ – set of acts

- $X=\{0, 10, 15, 20, 30, \ldots\}$ – set of possible outcomes (gains)

- $Cl=\{Cl_1, Cl_2, Cl_3, \ldots\}$ – set of quality classes of the acts,
  e.g.: $Cl_1=$bad acts, $Cl_2=$medium acts, $Cl_3=$good acts

- $\rho(A_i, \pi)=x$ means that by act $A_i$ one can gain at least $x$ with at least probability $\pi=Pr(W)$, where $W \subseteq ST$ is an event

- There is a partial preorder on probabilities $\pi$ of events

- Act $A_i$ stochastically dominates $A_j$ iff $\rho(A_i, \pi) \geq \rho(A_j, \pi)$
  for each probability $\pi$
DRSA for decision under risk and uncertainty

- Preference information given by a Decision Maker: assignment of some acts to quality classes

- Example:

<table>
<thead>
<tr>
<th>( \pi/Act )</th>
<th>( A_1 )</th>
<th>( A_2 )</th>
<th>( A_3 )</th>
<th>( A_4 )</th>
<th>( A_5 )</th>
<th>( A_6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.25</td>
<td>30</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>.35</td>
<td>10</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>.40</td>
<td>10</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>.60</td>
<td>10</td>
<td>20</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>.65</td>
<td>10</td>
<td>20</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>.75</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>15</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Class: good, medium, bad, medium, good
DRSA for decision under risk and uncertainty

- **Decision rules** induced from rough approximations of quality classes

\[
\text{if } \rho(A_i, 0.75) \geq 20 \text{ and } \rho(A_i, 1) \geq 10, \text{ then } A_i \in C \leq_3 \quad(A_6)
\]

“if the probability of gaining \textit{at least} 20 is \geq 0.75, and the probability of gaining \textit{at least} 10 is 1, then act \( A_i \) is \textit{at least good}”

\[
\text{if } \rho'(A_i, 0.25) \leq 20 \text{ and } \rho'(A_i, 0.75) \leq 15, \text{ then } A_i \in C \leq_2 \quad(A_3, A_4, A_5)
\]

“if the probability of gaining \textit{at most} 20 is \geq 0.25, and the probability of gaining \textit{at most} 15 is \geq 0.75, then act \( A_i \) is \textit{at most medium}”

- **Generalization:**
  DRSA for decision under risk with outcomes distributed over time
  (decision under uncertainty and time preference)

What other monotonic relationships can be handled by DRSA?

- **DRSA – dominance relation:**
  
  “The more, the better”

- **DRSA for decision under uncertainty – stochastic dominance relation:**
  
  “The more and the more probable, the better”

- **DRSA for time preferences – time dominance relation:**
  
  “The more and the earlier, the better”

- **DRSA for time preference & uncertainty – time stochastic dominance:**
  
  “The more, the more probable and the earlier, the better”

- **DRSA can be applied to a large collection of operational research problems**, such as portfolio selection, scheduling under uncertainty, inventory management, interactive (robust) multiobjective optimization, ...
Decision rule approach to interactive multiobjective optimization
DRSA to Interactive Multiobjective Optimization

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

subject to the constraints:

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

where \( x = [x_1, \ldots, x_k] \) is a vector of decision variables

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

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

\( b_i, i=1,\ldots,m, \) are constant RHS of the constraints
Multiobjective Optimization – dominance relation

- Solution \( a \in A \) is **Pareto-optimal** (non-dominated) if and only if there is no other solution \( b \in A \) such that \( f_i(b) \succeq f_i(a) \), \( i \in \{1, \ldots, n\} \), and on at least one objective \( j \in \{1, \ldots, n\} \), \( f_j(b) \succ f_j(a) \)

![Diagram of Pareto-optimal solutions]

\( f_1 \) and \( f_2 \) are to be minimized
Evolutionary Multiobjective Optimization (EMO)
Dominance-based association rules describing the Pareto optimal set

- Relationships between attainable values of different objective functions (criteria) in the set of Pareto-optimal solutions

- Formal syntax (in case of maximization of objectives):

  \[
  \text{If } f_{i1}(x) \geq r_{i1} \text{ and } f_{i2}(x) \geq r_{i2} \text{ and } \ldots \text{ and } f_{ip}(x) \geq r_{ip},
  \]

  \[
  \text{then } f_{ip+1}(x) \leq r_{ip+1} \text{ and } f_{ip+2}(x) \leq r_{ip+2} \text{ and } \ldots \text{ and } f_{iq}(x) \leq r_{iq}
  \]

- Example from product-mix problem:

  - “if profit \( \geq 148 \) & time\_\text{machine} \( \leq 150 \),

  \[
  \text{then } \text{amount\_product\_x}_B \leq 2
  \]

Multiobjective Optimization – interactive procedures
What preference information and preference model should be used?

- The traditional interactive methods appear to be too demanding of the cognitive effort of their users.
- We advocate for "easy" preference information = natural and partial.
- The most natural is a holistic comparison of some solutions.
- The preference model should be intelligible and comprehensible.
- We advocate for decision rules.
Interactive cycle with elicitation of preferences

- Decision Maker
- Preference information
- Inference engine
- Optimizer
- DRSA decision rules
- Preference model
- Set of solutions
- Holistic comparison of some solutions
DRSA to Interactive Multiobjective Optimization – DRSA-IMO

1) Present to DM a representative set of efficient (Pareto-optimal) solutions

2) Present association rules showing relationships between the attainable values of the objective functions and relationships between decision variables and objective functions in the Pareto-optimal set

3) If DM finds a satisfactory solution, then stop, otherwise go to step 4)

4) DM selects efficient solutions judged as (relatively) good and bad

5) DRSA „if..., then...“ decision rules are induced from info got in step 4)

6) The most interesting decision rules are presented to DM

7) The DM selects one or more decision rules being the most adequate to his/her preferences

8) Constraints relative to these decision rules are included in the set of constraints

9) Go back to step 1)
Examples of Applications of DRSA
The „at least“ rules

\[ \text{if } x_{q_1} \geq_{q_1} r_{q_1} \text{ and } x_{q_2} \geq_{q_2} r_{q_2} \text{ and } \ldots \text{ and } x_{q_p} \geq_{q_p} r_{q_p}, \text{ then } x \in \text{Class}_{t \geq} \]

indicate **opportunities for improving** the assignment of object \( x \) to \( \text{Class}_t \) or better, if it was not assigned as high, and its score on \( q_1, \ldots, q_p \) would grow to \( r_{q_1}, \ldots, r_{q_p} \)

The „at most“ rules

\[ \text{if } x_{q_1} \leq_{q_1} r_{q_1} \text{ and } x_{q_2} \leq_{q_2} r_{q_2} \text{ and } \ldots \text{ and } x_{q_p} \leq_{q_p} r_{q_p}, \text{ then } x \in \text{Class}_{t \leq} \]

indicate **threats for deteriorating** the assignment of object \( x \) to \( \text{Class}_t \) or worse, if it was not assigned as low, and its score on \( q_1, \ldots, q_p \) would drop to \( r_{q_1}, \ldots, r_{q_p} \)

\[
incr_{SS'}(\Psi) = \sum_{\emptyset \subseteq P \subseteq N} \left[ \text{cer}_S(\Phi, \Psi) \times \text{cer}_{S'}(\neg \Psi, \bigwedge_{i \in P} \neg \Phi_i \wedge \bigwedge_{j \notin P} \Phi_j) \right] \times \frac{\|\neg \Psi\|_{S'}}{|U'|} \]

Intervention based on „at least” and „at most” rules - example

- Example: customer satisfaction analysis by a Company
- 44 questions and 3 classes of overall satisfaction: High, Medium, Low

Threats of deterioration of satisfaction

- Deterioration from High or Medium to Low satisfaction

Opportunities for improvement of satisfaction

- Improvement from Low to Medium or High satisfaction
- Improvement from Low or Medium to High satisfaction

Deterioration from High to Medium or Low satisfaction
Factors for Consumer Channel

A: Product and product quality
- A1: Width of product range
- A2: Fulfilling Account’s shopper’s/the end consumer’s needs
- A3: Innovative product introducing
- A4: Quality of supplier’s products from the consumer perspective
- A5: Price/quality ratio of the light bulbs from the consumer perspective
- A6: Packaging

B: Logistics
- B1: Quality of changes implementing process
- B2: Masterdata
- B3: Availability of ordered products
- B4: Accuracy of the delivery
- B5: The punctuality of the date of the delivery
- B6: Communication on delivery date changes
- B7: Condition of delivered products
- B8: Delivery documents

C: Customer service
- C1: Accessibility of KAM
- C2: KAM’s knowledge about products
- C3: Professionalism of KAM
- C4: Speed of response to problems
- C5: Ease of order placing
- C6: Speed of order processing
- C7: Satisfaction with settlement of complaints
- C8: Timely credit/debit payments
- C9: Clarity of invoices
- C10: Accuracy of invoices
- C11: Logic of shelf display
- C12: Stock replenishment adjusted to the speed of product rotation
- C13: Order/neatness of the shelf
- C14: Service merchandising costs

D: Marketing support
- D1: Sufficient number of promotions offered by the supplier
- D2: Contribution to Account’s sales increase
- D3: Contribution to Account’s strategy for the category
- D4: Availability of POS materials
- D5: Quality of POS materials
- D6: MarCom support during product introductions
- D7: Shoppers shelf perception
- D8: Optimizing the profit of the shelf
- D9: Satisfying Account’s strategy
- D10: Possibility and effectiveness of training
- D11: Effectiveness of product information

F: Business relationship with the supplier
- F1: Profit Margins
- F2: Product price / performance ratio
- F3: Focus on customer service
- F4: Focus on end consumers
- F5: Supplier’s market leadership

44 factors organized into 5 groups (A-D, F)
Decision rules induced from customer data structured by DRSA

Certain **at least** rules

- If \((F_1 \geq 5) \Rightarrow (SATISFACTION \geq \text{HIGH})\)
- If \((A_1 \geq 4) \& (E_2 \geq 5) \Rightarrow (SATISFACTION \geq \text{HIGH})\)
- If \((A_3 \geq 5) \& (C_3 \geq 5) \Rightarrow (SATISFACTION \geq \text{HIGH})\)
- If \((A_1 \geq 5) \& (C_4 \geq 5) \Rightarrow (SATISFACTION \geq \text{HIGH})\)
- If \((F_1 \geq 4) \Rightarrow (SATISFACTION \geq \text{MEDIUM})\)
- If \((A_1 \geq 4) \& (C_3 \geq 3) \Rightarrow (SATISFACTION \geq \text{MEDIUM})\)

Certain **at most** rules

- If \((C_4 \leq 2) \Rightarrow (SATISFACTION \leq \text{LOW})\)
- If \((F_1 \leq 2) \Rightarrow (SATISFACTION \leq \text{LOW})\)
- If \((A_1 \leq 2) \Rightarrow (SATISFACTION \leq \text{MEDIUM})\)
- If \((C_1 \leq 2) \Rightarrow (SATISFACTION \leq \text{MEDIUM})\)
- If \((B_2 \leq 2) \Rightarrow (SATISFACTION \leq \text{MEDIUM})\)
- If \((E_3 \leq 3) \Rightarrow (SATISFACTION \leq \text{MEDIUM})\)
- If \((A_3 \leq 4) \& (A_4 \leq 4) \Rightarrow (SATISFACTION \leq \text{MEDIUM})\)
- If \((A_3 \leq 4) \& (C_3 \leq 4) \Rightarrow (SATISFACTION \leq \text{MEDIUM})\)
Intervention based on monotonic rules - example

*At least* rule:

**If** \((A3 \geq 4) \& (C3 \geq 3)\), **then** Satisfaction \(\succ Medium\)

\[ \text{incr}_{SS}(Medium) = 77\% \]

*Opportunity*: if

- \(A3 \geq 4\), and
- \(C3 \geq 3\), then

satisfaction of 77% of customers with \(\text{Satisfaction} = Low\) will **improve** to **Medium** or **High**
At most rule:

If \((A2 \leq 3) \& (E4 \leq 4)\), then Satisfaction \(\leq Low\)

\(\text{incr}_{SS}'(Low) = 89\%\)

Threat: if

- \(A2 \leq 3\), and
- \(E4 \leq 4\), then

satisfaction of 89\% of customers with Satisfaction = High or Medium will deteriorate to Low
Intervention based on monotonic rules

- In practice, the choice of rules used for intervention can be supported by additional measures, like:
  - length of the rule – the shorter the better,
  - cost of intervention on attributes present in the rule,
  - priority of intervention on some types of attributes, like: short-term before long-term actions
Mobile Emergency Triage System - MET System

- MET – Mobile Emergency Triage
  - Facilitates triage disposition for presentations of acute pain (abdominal and scrotal pain, hip pain)
  - Supports triage decision with or without complete clinical information
  - Provides mobile support through handheld devices
  - http://www.mobiledss.uottawa.ca

W. Michalowski, University of Ottawa
K. Farion, Children’s Hospital of Eastern Ontario
Sz. Wilk, R. Słowiński, Poznań University of Technology
Triage Process

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

Emergency Room (ER)

Hospital/Clinic

- Observation
- Discharge

Management

Diagnosis and treatment

Triage
### Decision Rules (examples)

- **if** (Age < 5 years) **and** (PainSite = lower_abdomen) **and** (RebTend = yes) **and** (4 < WBC < 12) **then** (Triage = discharge)

- **if** (PainDur > 7 days) **and** (PainSite = lower_abdomen) **and** (37 ≤ Tempr ≤ 39) **and** (TendSite = lower_abdomen) **then** (Triage = observation)

- **if** (Sex = male) **and** (PainSite = lower_abdomen) **and** (PainType = constant) **and** (RebTend = yes) **and** (WBCC ≥ 12) **then** (Triage = consult)
System MET-AP
Violinmakers competition

Criteria:
- volume of sound (X),
- timbre of sound (Y),
- ease of sound emission,
- equal sound volume of strings (Z),
- accuracy of assembly,
- individual qualities

Jury’s assessment

Sound recording

The violin’s acoustic data:
- individual sounds played on open strings, G,D,A,E,
- successive sounds of chromatic scale,

Acoustic features:
- power spectrum of chromatic scale sounds,
- wavelets,
- harmonic based spectral parameters (tristimuli, brightness, odd/even harmonics content...),
- psychoacoustic features
- cepstral coefficients.

Dominance-based Rough Set Approach
Violinmakers competition – DRSA results

- Reconstructing the expert’s rankings of a set of 23 violins
- Three rankings: volume, timbre and inter-string equality
- Feature space - cepstral coefficients

<table>
<thead>
<tr>
<th>Ranking according to</th>
<th>Best subset of acoustic features</th>
<th>Number of rules</th>
<th>Ranking fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>volume</td>
<td>A14, E13, D12, G16</td>
<td>62</td>
<td>87%</td>
</tr>
<tr>
<td>timbre</td>
<td>E13, D15, G4, G17, D5</td>
<td>99</td>
<td>92%</td>
</tr>
<tr>
<td>inter-string equality</td>
<td>D20, D15, A24, D10</td>
<td>64</td>
<td>79%</td>
</tr>
</tbody>
</table>
Technical diagnostics – problem of short circuits in coil body

- An element of a light-bulb
Technical diagnostics – problem of short circuits in coil body

- Problem – coil geometry failure – short circuits in coil body
Technical diagnostics – problem of short circuits in coil body

- Production process steps

1st Coiling
Potflyer Mo mandrel

2nd Coiling
BH Mo mandrel

Annealing
Tandem

Cutting
SAM

Sintering
NT Furnace

Dissolving
Mo
Technical diagnostics – problem of short circuits in coil body

- **Issues:**
  - Wire diameter ($W \sim 20\mu m; Mo \sim 175\mu m$)
  - Batch throughput time (avg. 10 days)
  - Many factors suspected for failure, including interactions
    - Materials
    - Subprocesses
    - People
  - Such coil geometry failure occurred first time in history
  - Defects are hardly visible on machines
Technical diagnostics – problem of short circuits in coil body

- **Data table** – 550 lots described by 10 attributes

<table>
<thead>
<tr>
<th>X</th>
<th>B</th>
<th>G</th>
<th>I</th>
<th>J</th>
<th>L</th>
<th>N</th>
<th>P</th>
<th>V</th>
<th>W</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Month &amp; day of the month</td>
<td>ID of wolfram lot</td>
<td>ID of 1st coiling machine</td>
<td>No. of breaks</td>
<td>ID of 2nd coiling machine</td>
<td>ID of annealing furnace</td>
<td>ID of cutting machine</td>
<td>Day of the week</td>
<td>Number of days in production</td>
<td>Failure</td>
</tr>
<tr>
<td>2</td>
<td>11_08</td>
<td>M296</td>
<td>26/3</td>
<td>42</td>
<td>75/3</td>
<td>5</td>
<td>6</td>
<td>Mo</td>
<td>15</td>
<td>NO</td>
</tr>
<tr>
<td>3</td>
<td>11_02</td>
<td>M297</td>
<td>7/2</td>
<td>37</td>
<td>6/1</td>
<td>5</td>
<td>3</td>
<td>Fr</td>
<td>34</td>
<td>YES</td>
</tr>
<tr>
<td>4</td>
<td>10_15</td>
<td>M288</td>
<td>7/4</td>
<td>13</td>
<td>7/1</td>
<td>5</td>
<td>6</td>
<td>Tu</td>
<td>21</td>
<td>NO</td>
</tr>
<tr>
<td>5</td>
<td>11_09</td>
<td>M296</td>
<td>48/6</td>
<td>0</td>
<td>6/1</td>
<td>5</td>
<td>1</td>
<td>Mo</td>
<td>32</td>
<td>NO</td>
</tr>
<tr>
<td>6</td>
<td>11_21</td>
<td>M286</td>
<td>26/6</td>
<td>14</td>
<td>7/4</td>
<td>5</td>
<td>1</td>
<td>We</td>
<td>33</td>
<td>NO</td>
</tr>
<tr>
<td>7</td>
<td>11_37</td>
<td>M293</td>
<td>26/4</td>
<td>12</td>
<td>75/3</td>
<td>5</td>
<td>1</td>
<td>Tu</td>
<td>21</td>
<td>NO</td>
</tr>
<tr>
<td>8</td>
<td>11_04</td>
<td>M292</td>
<td>7/1</td>
<td>10</td>
<td>75/3</td>
<td>5</td>
<td>1</td>
<td>Fr</td>
<td>21</td>
<td>YES</td>
</tr>
<tr>
<td>9</td>
<td>11_43</td>
<td>M296</td>
<td>7/2</td>
<td>16</td>
<td>75/4</td>
<td>5</td>
<td>1</td>
<td>Th</td>
<td>13</td>
<td>NO</td>
</tr>
<tr>
<td>10</td>
<td>11_20</td>
<td>M296</td>
<td>48/3</td>
<td>0</td>
<td>34/4</td>
<td>12</td>
<td>3</td>
<td>Mo</td>
<td>35</td>
<td>YES</td>
</tr>
<tr>
<td>11</td>
<td>11_10</td>
<td>M286</td>
<td>7/3</td>
<td>0</td>
<td>70/2</td>
<td>5</td>
<td>3</td>
<td>We</td>
<td>35</td>
<td>NO</td>
</tr>
<tr>
<td>12</td>
<td>11_17</td>
<td>M298</td>
<td>7/4</td>
<td>18</td>
<td>4/1</td>
<td>12</td>
<td>3</td>
<td>Tu</td>
<td>43</td>
<td>YES</td>
</tr>
<tr>
<td>13</td>
<td>11_07</td>
<td>M296</td>
<td>7/3</td>
<td>0</td>
<td>73/4</td>
<td>5</td>
<td>3</td>
<td>Fr</td>
<td>30</td>
<td>YES</td>
</tr>
<tr>
<td>14</td>
<td>11_34</td>
<td>M293</td>
<td>23/1</td>
<td>0</td>
<td>6/3</td>
<td>5</td>
<td>1</td>
<td>Th</td>
<td>20</td>
<td>NO</td>
</tr>
<tr>
<td>15</td>
<td>10_44</td>
<td>M288</td>
<td>48/5</td>
<td>0</td>
<td>70/1</td>
<td>5</td>
<td>1</td>
<td>We</td>
<td>29</td>
<td>NO</td>
</tr>
<tr>
<td>16</td>
<td>11_01</td>
<td>M288</td>
<td>7/3</td>
<td>28</td>
<td>3/4</td>
<td>1</td>
<td>3</td>
<td>Tu</td>
<td>0</td>
<td>NO</td>
</tr>
<tr>
<td>17</td>
<td>11_12</td>
<td>M296</td>
<td>7/2</td>
<td>14</td>
<td>73/1</td>
<td>12</td>
<td>1</td>
<td>Mo</td>
<td>67</td>
<td>YES</td>
</tr>
<tr>
<td>18</td>
<td>11_20</td>
<td>M300</td>
<td>48/6</td>
<td>0</td>
<td>33/3</td>
<td>9</td>
<td>6</td>
<td>Tu</td>
<td>33</td>
<td>NO</td>
</tr>
<tr>
<td>19</td>
<td>10_22</td>
<td>M287</td>
<td>48/4</td>
<td>0</td>
<td>71/4</td>
<td>6</td>
<td>6</td>
<td>We</td>
<td>26</td>
<td>YES</td>
</tr>
<tr>
<td>20</td>
<td>11_18</td>
<td>M299</td>
<td>7/4</td>
<td>17</td>
<td>36/4</td>
<td>2</td>
<td>6</td>
<td>Th</td>
<td>33</td>
<td>NO</td>
</tr>
<tr>
<td>21</td>
<td>11_44</td>
<td>M296</td>
<td>48/1</td>
<td>0</td>
<td>6/4</td>
<td>9</td>
<td>1</td>
<td>Fr</td>
<td>14</td>
<td>YES</td>
</tr>
<tr>
<td>22</td>
<td>11_06</td>
<td>M299</td>
<td>7/5</td>
<td>20</td>
<td>20</td>
<td>1</td>
<td>3</td>
<td>Mo</td>
<td>35</td>
<td>NO</td>
</tr>
</tbody>
</table>
Technical diagnostics – problem of short circuits in coil body

- Results of the DRSA analysis:
- Quality of approximation of the classification into Yes/No failure: 100%
- Reducts: 61 with 8 to 5 attributes
- Adopted reduct (5 attributes):
  - I – ID of the 1st coiling machine
  - L – ID of the 2nd coiling machine
  - N – ID of the annealing furnace
  - B – lot ID: day of the month
  - V – lot ready: day of the week
Technical diagnostics – problem of short circuits in coil body

- Rules:

If 1st coiling machine = 13, then YES failure
   support = 8%

If 1st coiling machine = 48 & Furnace = 12 & Cutting machine = 3,
   then YES failure
   support = 12%

If Furnace = 12 & Cutting machine = 3 & Day = Friday, then YES failure
   support = 9%

If Furnace = 5 & Cutting machine = 6, then NO failure
   support = 14.67%

If Furnace = 5 & Month of the lot = 12, then NO failure
   support = 20.22%
Other applications of DRSA

- Prediction of Antimicrobial Activity of Quaternary Ammonium Chlorides by analysis of structure-activity relationship (SAR)
- Other medical applications:
  - Complications after open-heart operations
  - Colon cancer surgery
  - Pediatric hip surgery
  - Prostate cancer treatment
  - Breast cancer treatment
  - HSV treatment of duodenal ulcer
  - Extracorporeal shockwave lithotripsy (ESWL)
  - Asthma treatment
  - ...
Other methodological extensions of DRSA

- DRSA for Choice and Ranking with multi-graded preference relations
- DRSA as a Way of Handling Fuzzy-Rough Hybridization
- DRSA for Case-Based Reasoning
- DRSA for Decision Under Uncertainty and Time Preference
- DRSA for Ordinal Classification with Imprecise or Missing Evaluations and Assignments
- DRSA for Group Decision
- DRSA for Hierarchical Structure of Attributes and Criteria
- DRSA for Financial Portfolio Decision
- DRSA for Customer Satisfaction Analysis
Conclusions

- Monotonic "if..., then..." decision rules give account of most complex interactions among attributes, require weaker axioms than other preference models, and can represent inconsistent preferences.

- Heterogeneous information (attributes, criteria) and attribute scales (ordinal, cardinal) can be handled by DRSA.

- DRSA exploits ordinal information only, and decision rules do not convert ordinal information into numeric one.

- DRSA supplies useful elements of knowledge about decision situation:
  - certain and doubtful knowledge distinguished by lower and upper appx.
  - relevance of particular attributes and information about their interaction,
  - reducts & core of attributes conveying important knowledge contained in data,
  - decision rules can be used for explanation of past decisions, for decision support and for strategic interventions.

- DRSA has sound theoretical foundations (bipolar algebra, bitopology, Bayesian confirmation theory)
Software available on the web

**ROSE**
ROugh Set data Explorer

**4eMka & JAMM & jMAF**
New Decision Support Tools for Rule-based Analysis and Solving of Multi-attribute Decision Problems


THANK YOU!