Polish Academy of Sciences

Constructive Preference Learning

robust ordinal regression for multi-attribute decision aiding

Roman Słowiński

Poznań University of Technology & Polish Academy of Sciences, Poland

International Conference on Operations Research and Enterprise Systems (ICORES), February 4-6, 2021

Plan

- Introduction – where is the challenge?
- Robust ordinal regression for value function preference model
  - Representative instance of the preference model
  - Extreme ranking analysis
  - Stochastic ordinal regression
  - Robust ordinal regression for hierarchy of criteria
  - Robust ordinal regression for group decision
- Robust ordinal regression for outranking relation preference model
- Robust ordinal regression for decision rule preference model
- Evolutionary Multiobjective Optimization involving ROR
- Decision under Risk and Uncertainty involving ROR
- Preference elicitation and justification of recommendations
- Summary and conclusions

Introduction – where is the challenge?

Decision problem

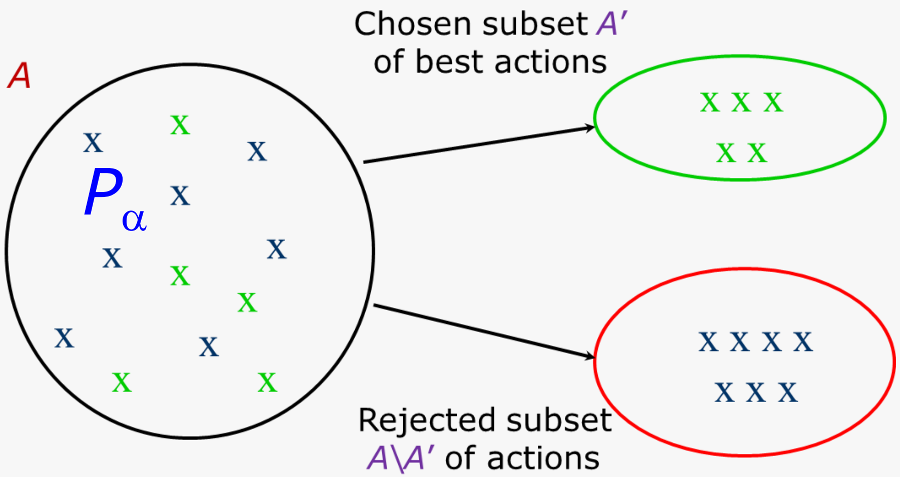
- There is an objective or objectives to be attained
- There are many alternative ways for attaining the objective(s) – they consititute a set of actions *A* (alternatives, solutions, objects, acts, …)
- Questions with respect to set *A*:

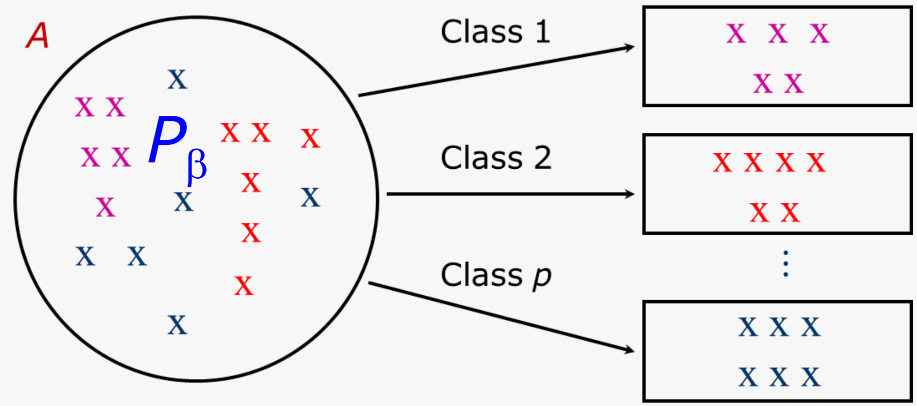
  P_α: *How to choose the best action* ?

  P_β : *How to classify actions into pre-defined decision classes* ?

  P_γ : *How to order actions from the best to the worst* ?

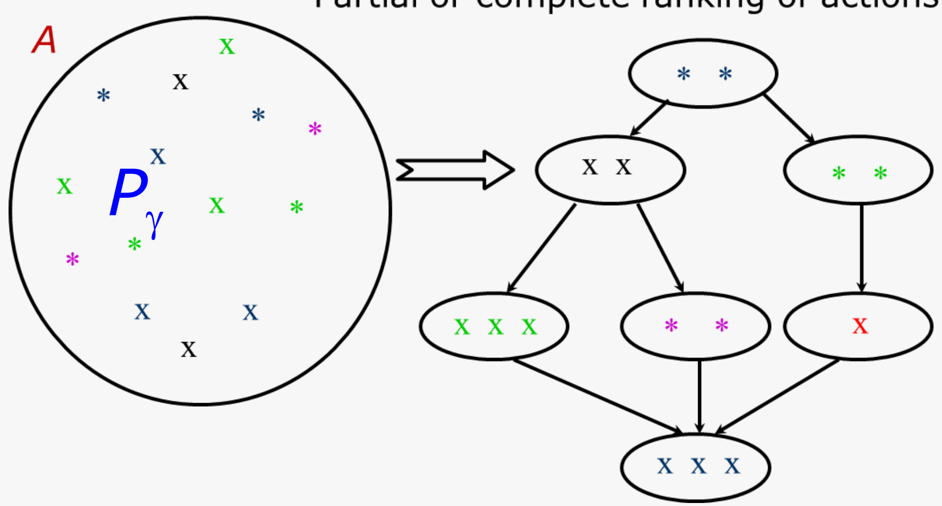
Decision problem





Class 1 $\succ$ Class 2 $\succ$ … $\succ$ Class p

Partial or complete ranking of actions



Coping with multiple dimensions in Decision Aiding

- Decision problems P_α, P_β , P_γ 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)

S. Greco, M. Ehrgott, J. Figueira (eds.), *Multiple Criteria Decision Analysis: State of the Art Surveys*. 2nd edition, OR & MS 233, Springer, New York, 2016

S. Greco, M. Ehrgott, J. Figueira (eds.), *Trends in Multiple Criteria Decision Analysis*. Springer, New York, 2010

Multi-dimensional decision problems

	Social Choice (Group Decision)	Multiple Criteria Decision Aiding	Decision under Risk and Uncertainty
Element of set *A*	Candidate	Action	Act
Dimension of evaluation space	Voter	Criterion	Probability of an outcome
Objective information about comparison of elements from *A*	Dominance relation	Dominance relation	Stochastic dominance relation

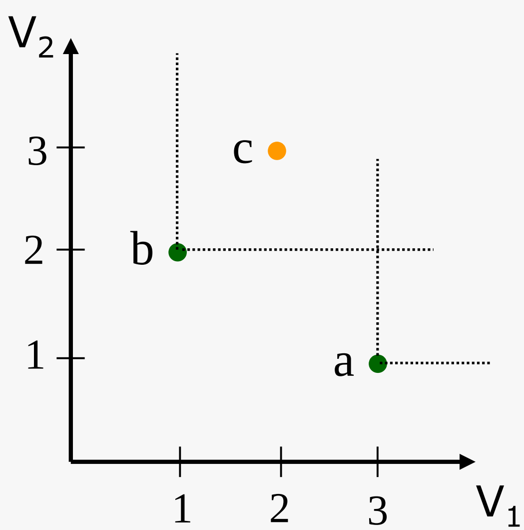
- The only objective information one can draw from the statement of a multi-dimensional decision problem is the **dominance relation**

SC&GD

Cand.	Voters: V_1	Voters: V_2
a	3	1
b	1	2
c	2	3

$V_1 : b \succ c \succ a$

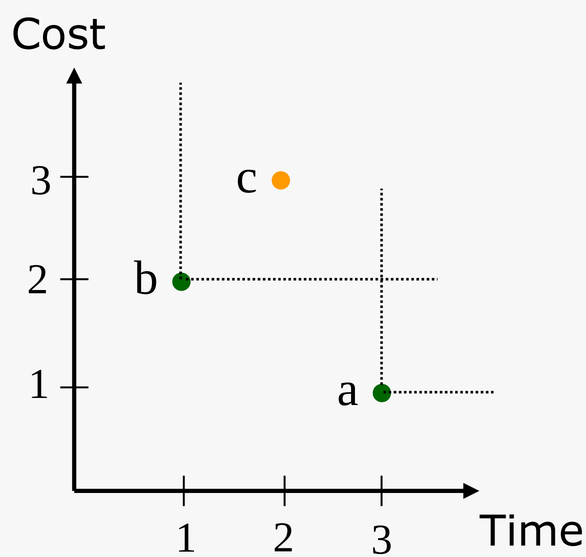
$V_2 : a \succ b \succ c$



MCDA

Action	Criteria: Time	Criteria: Cost
a	3	1
b	1	2
c	2	3

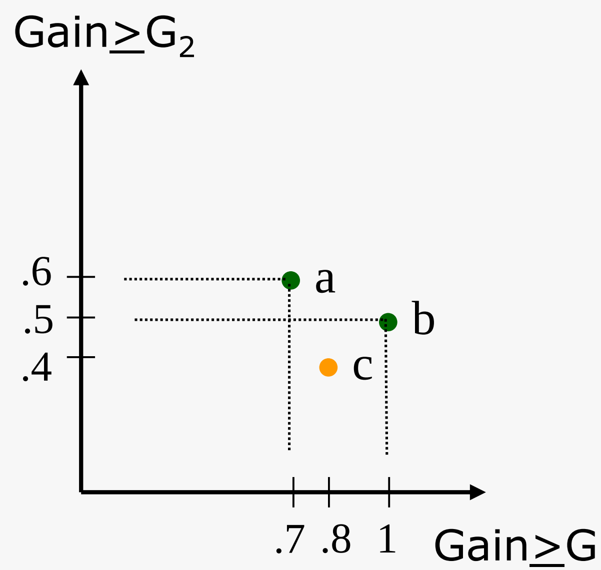
- non-dominated
- dominated



DRU

Act	Probability of gain: Gain$\geq G_1$	Probability of gain: Gain$\geq G_2$
a	0.7	0.6
b	1.0	0.5
c	0.8	0.4

$G_1 < G_2$



Enriching dominance relation – preference modeling/learning

- Dominance relation is too poor – it leaves many actions non-comparable

1

- One can „enrich" the dominance relation, using preference information elicited from the DM

2

- 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)

3

- 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 ranking

Aggregation of multiple criteria evaluations – preference models

- Three families of **preference modeling (aggregation) methods**:
  - Multiple Attribute Utility Theory (MAUT) using a value function,

    e.g., $U(a)=\sum_{i=1}^{n} w_i g_i(a)$, $U(a)=\sum_{i=1}^{n} u_i[g_i(a)]$, Choquet/Sugeno integral

  - Outranking methods using an outranking relation $S=\{\sim \cup \succ^{w} \cup \succ^{s}\}$

    $a\ S\ b$ = „a is at least as good as b"

  - Decision rule approach using a set of decision rules

e.g., „If $g_i(a)\succeq r_i$ & $g_j(a)\succeq r_j$ & ... $g_h(a)\succeq r_h$, then $a \rightarrow$ *Class* t or higher"

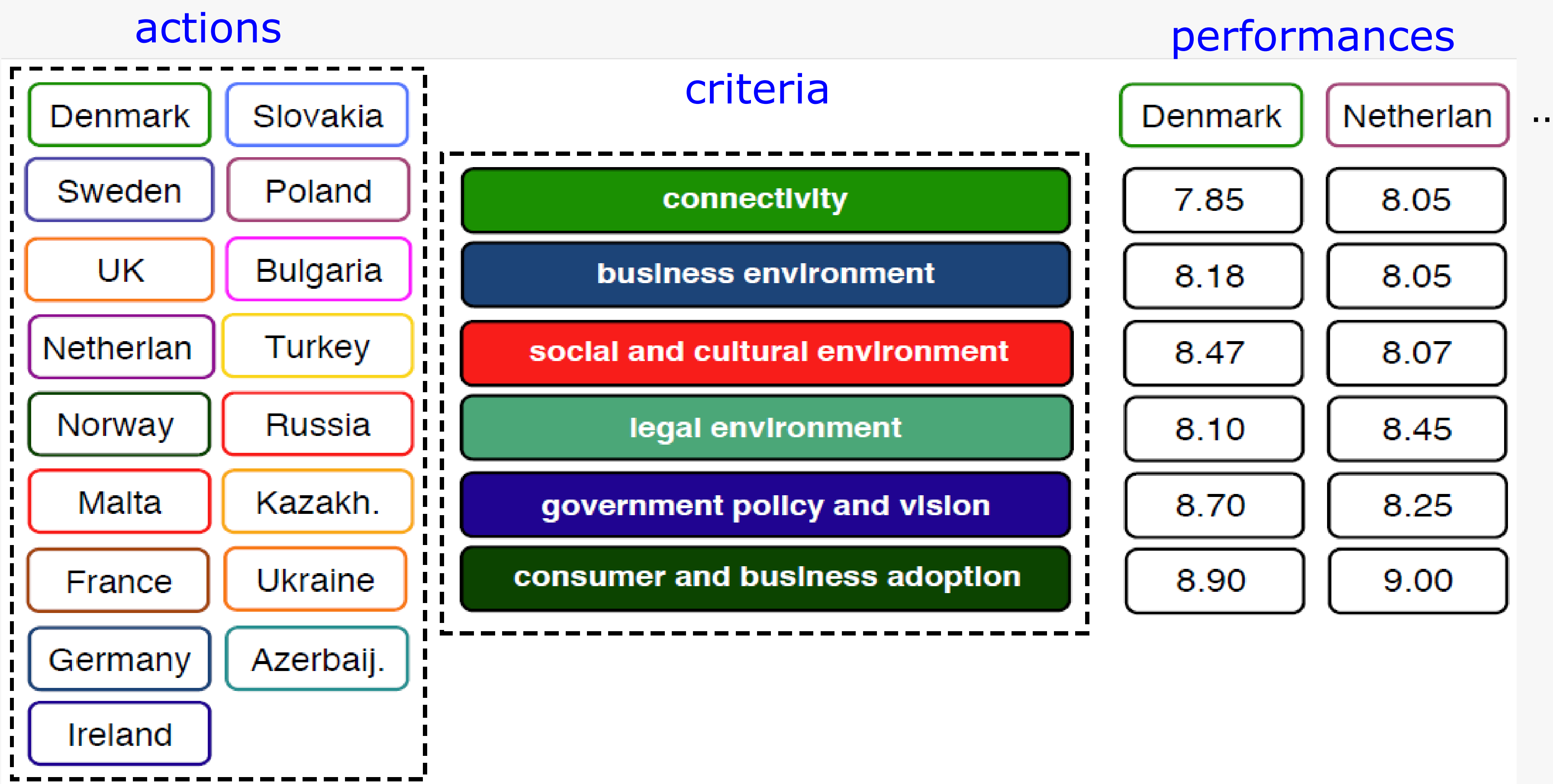
„If $g_i(a)\succeq_i^{\geq h(i)} g_i(b)$ & $g_j(a)\succeq_j^{\geq h(j)} g_j(b)$ & ... $g_p(a)\succeq_p^{\geq h(p)} g_p(b)$, *then* aSb"

  - Decision rule model is the most general of all three

R. Słowiński, S. Greco, B. Matarazzo: Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification problems under partial inconsistency with the dominance principle, *Control & Cybernetics*, 31 (2002) no.4, 1005-1035

Example

- Ranking of countries wrt digital economy (quality of information and technology infrastructure) (Economist Intelligence Unit in 2010)



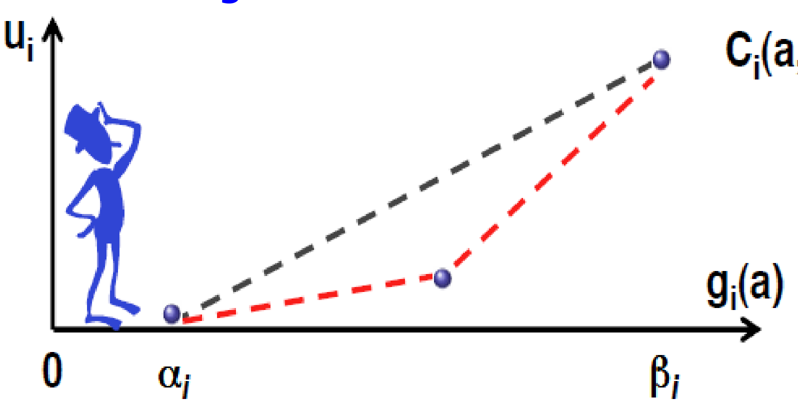
Elicitation of preference information by the Decision Maker (DM)

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

P.C.Fishburn (1967): Methods of Estimating Additive Utilities. *Management Science*, 13(7), 435-453 (listed and classified twenty-four methods of estimating additive utilities)

Value function model

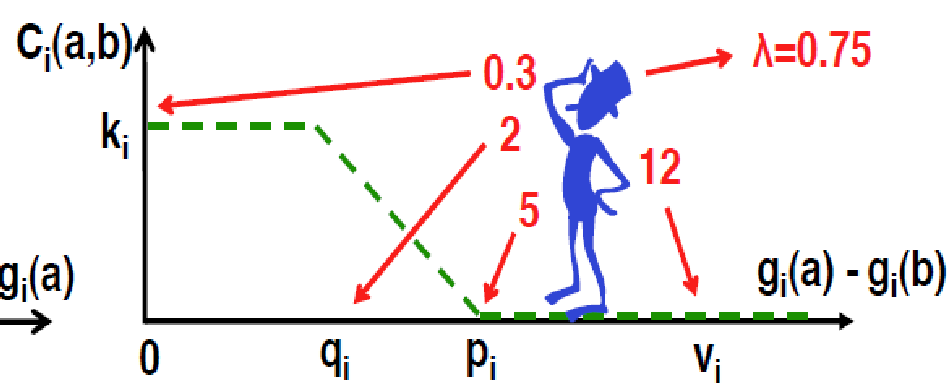
substitution rates or shapes of marginal value functions



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

Outranking model

weights & discrimination thresholds



$$aSb \Leftrightarrow C(a,b)=\sum_{i=1}^{n} C_i(a,b) \geq \lambda$$

and $g_i(b)-g_i(a) \leq v_i$ 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

[TIME=24, COST=56, RISK=75]

$\succ$

[TIME=28, COST=67, RISK=25]

Pairwise preferences between alternatives

$\succ$

$\succ$

characterized by cardinal and/or ordinal features (criteria)

[MATH=18, PHYS=16, LIT=15] $\Rightarrow$ Class „MEDIUM”
[MATH=17, PHYS=16, LIT=18] $\Rightarrow$ Class „GOOD”

Classification examples

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

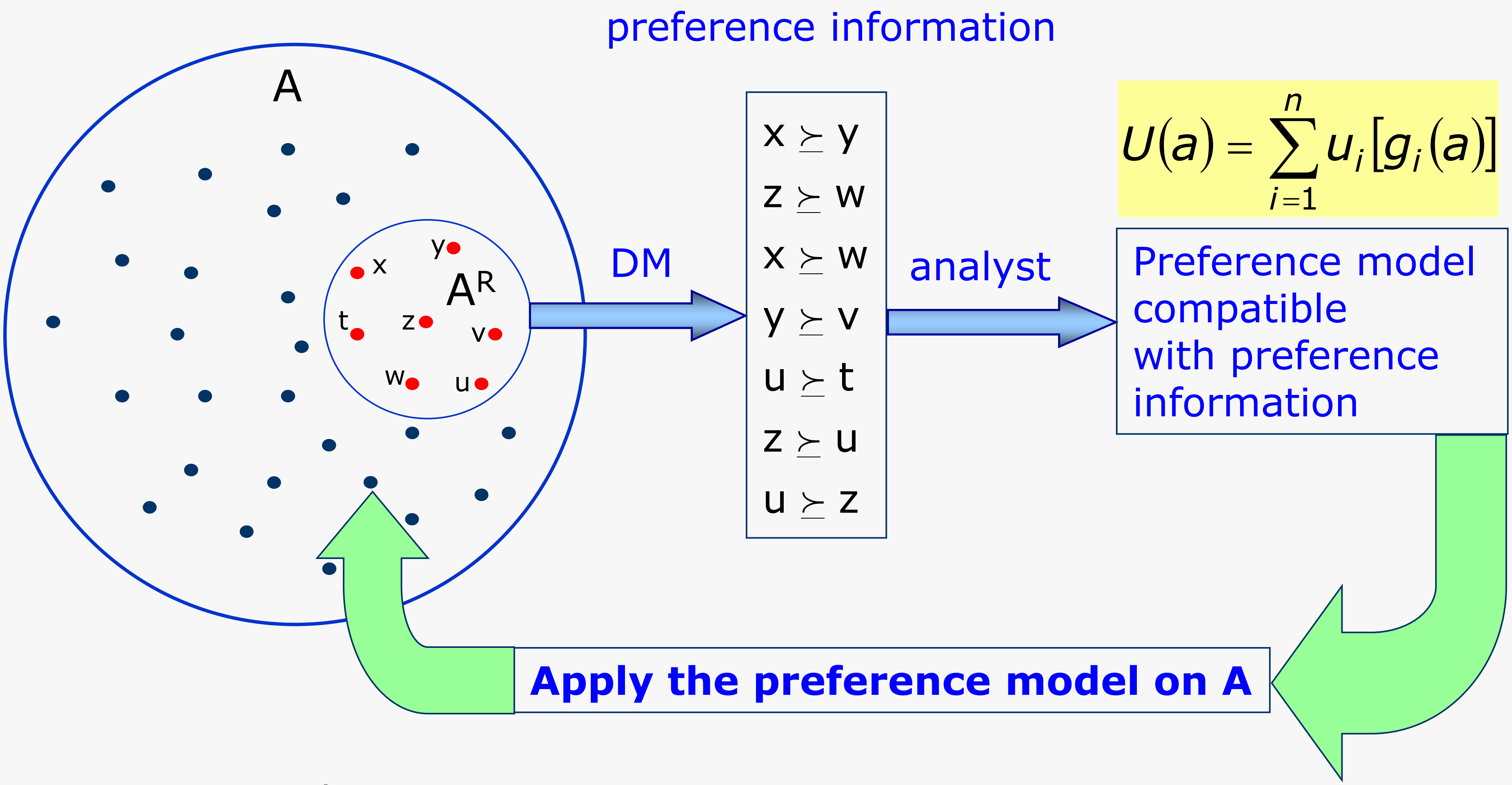
Intensity of preference

Alternative **F** should be among **5%** of the best ones

Rank related

Ordinal regression paradigm (UTA method)

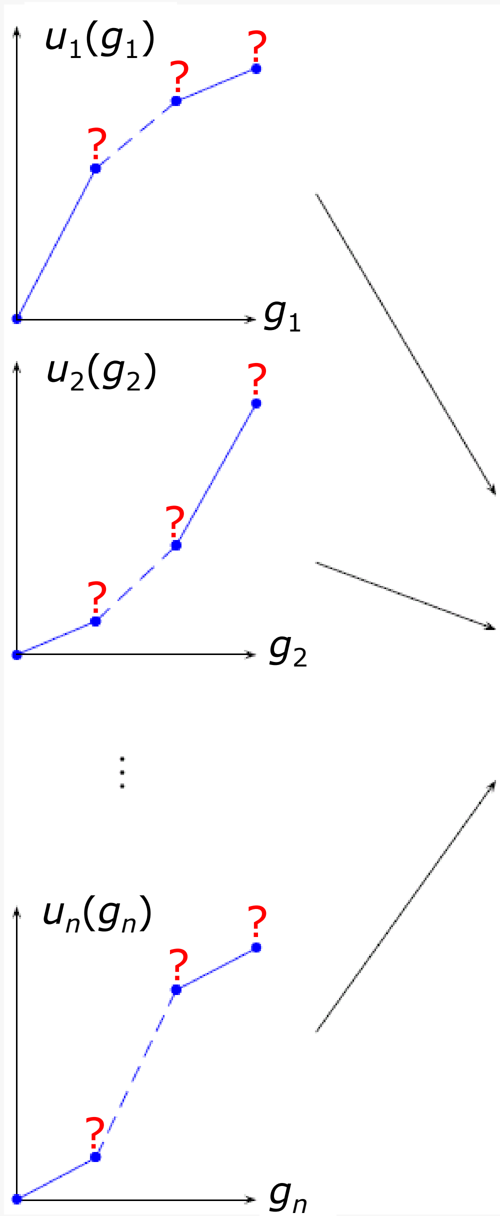
- Ordinal regression paradigm emphasizes the discovery of intentions as an interpretation of decision examples rather than as position *a priori*



E. Jacquet-Lagrèze, J. Siskos: Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. *Europ. J. Operational Research*, 10 (1982) 151-164

UTA additive preference model

Marginal value functions $u_i(g_i)$



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

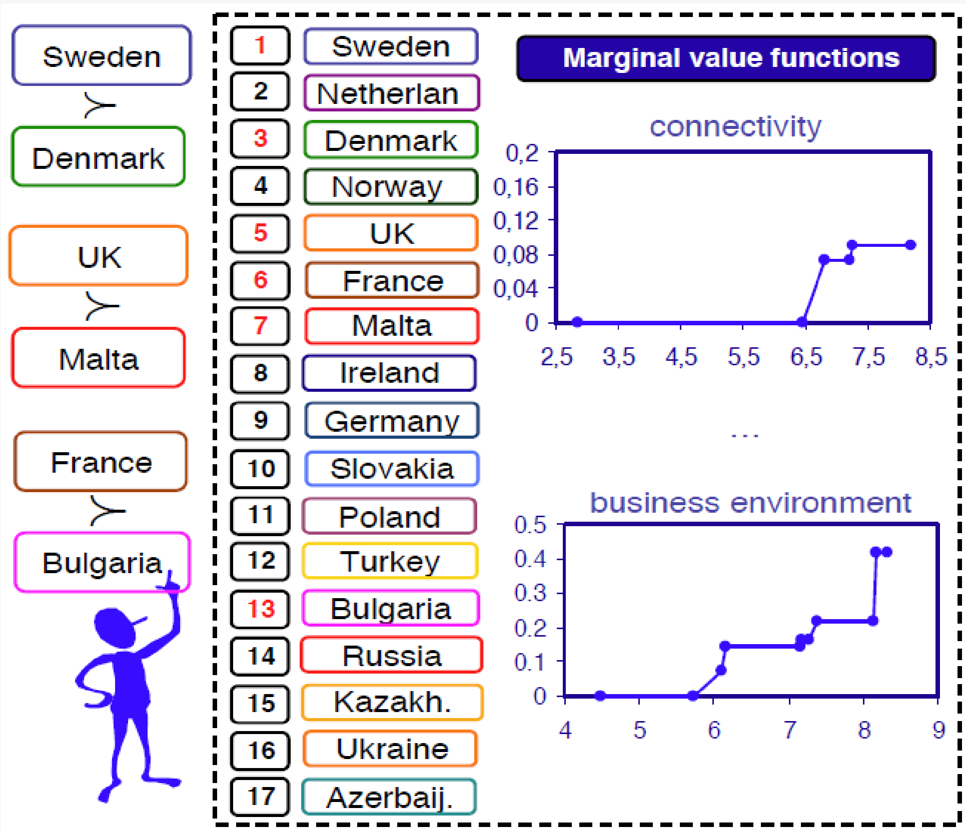
having evaluations $g_i(x),\ i = 1,\ldots,n$

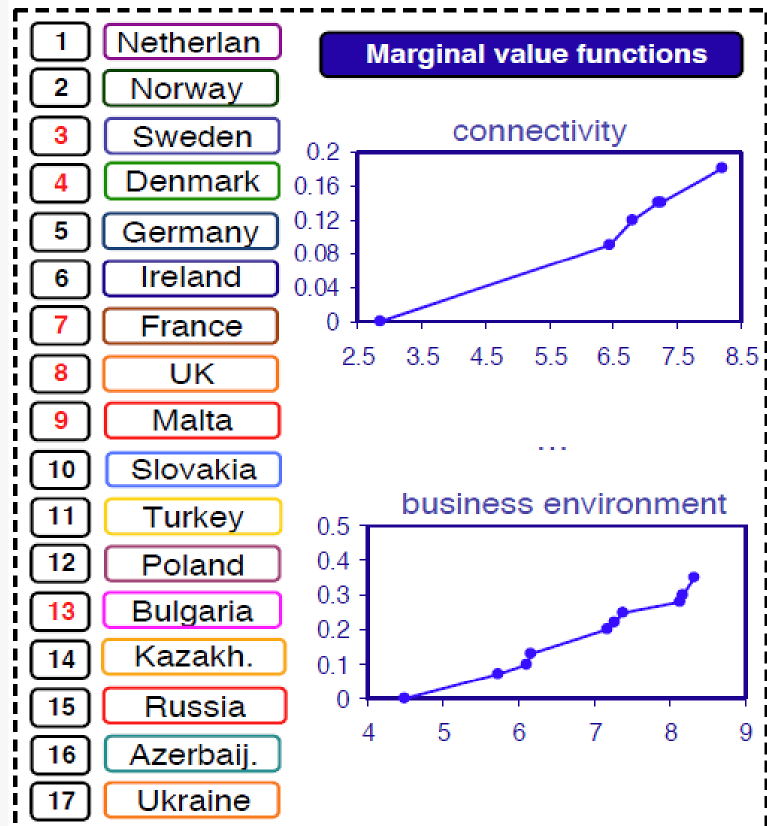
Criteria are supposed to be independent with respect to preferences

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

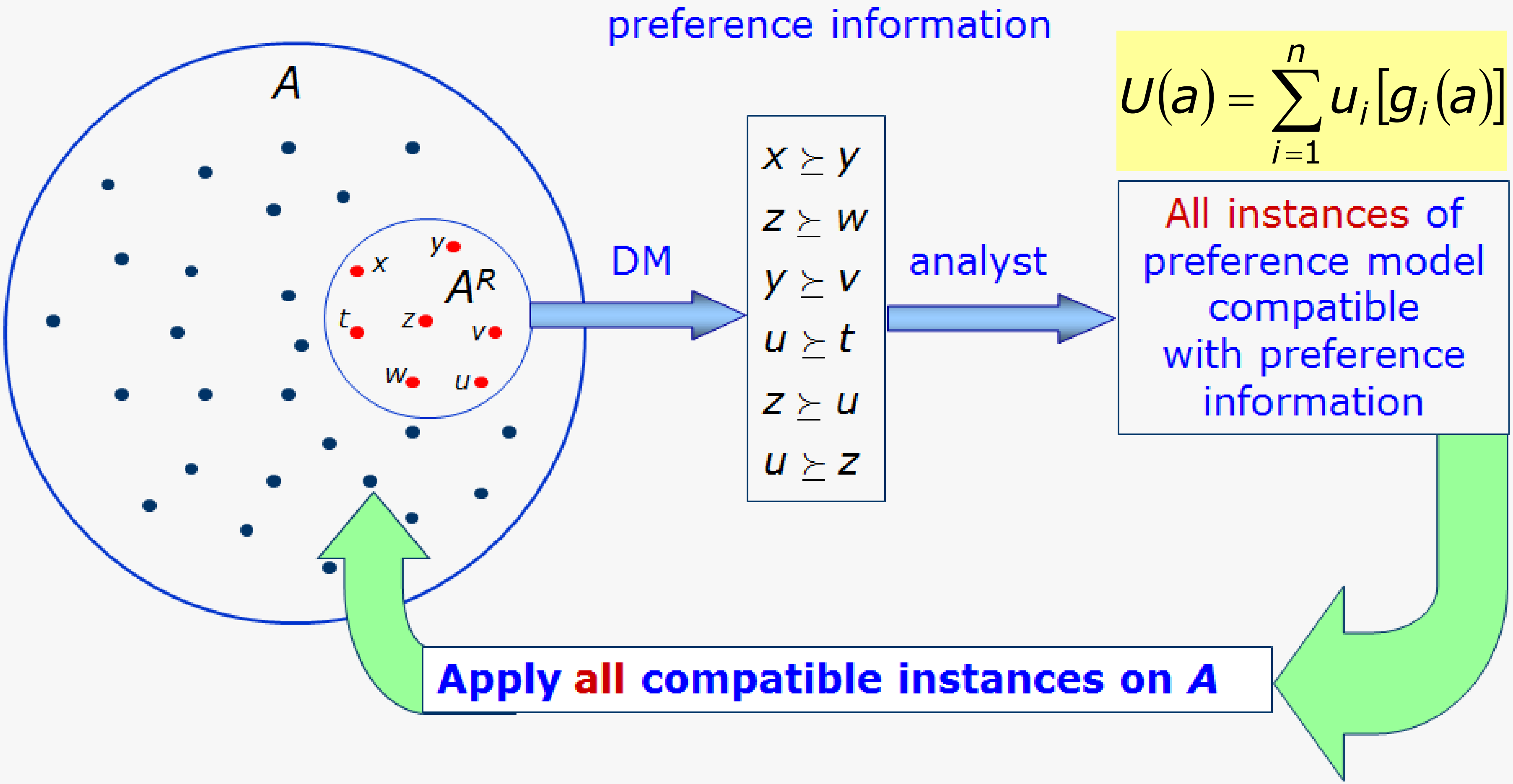




The two rankings are substantially different,
although both reproduce the same preference information

Robust Ordinal Regression
for value function preference model

Non-univocal representation - **R**obust **O**rdinal **R**egression - UTA^{GMS}



S. Greco, V. Mousseau, R. Słowiński: Ordinal regression revisited: multiple criteria ranking with a set of additive value functions. *European J. Operational Research*, 191 (2008) 415-435

ROR – possible and necessary preference relations

- The **possible** preference relation: for all alternatives $x,y \in A$,

  $x \succsim^P y \Leftrightarrow U(x) \geq U(y)$ **for at least one** compatible value function

  (complete and negatively transitive)

- The **necessary** preference relation: for all alternatives $x,y \in A$,

  $x \succsim^N y \Leftrightarrow U(x) \geq U(y)$ **for all** compatible value functions

  (partial preorder)

- When there is no preference information:

  necessary relation = dominance relation

$x \succsim^N y \;\Rightarrow\; x \succsim^P y$,

i.e., $\succsim^N \subseteq \succsim^P$

$x \succsim^N y$ or $y \succsim^P x$

for all $x,y \in A$

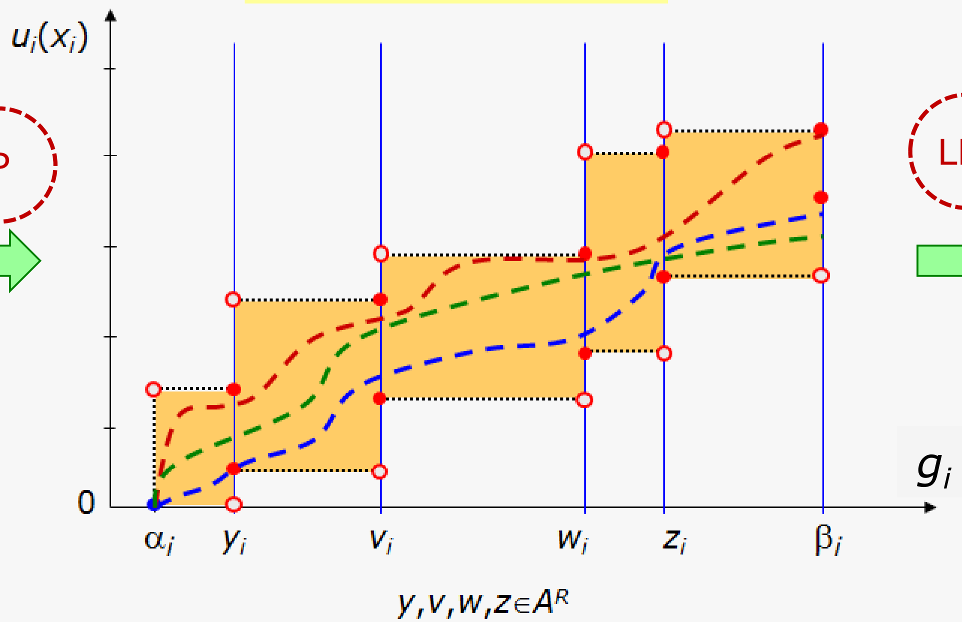
Non-univocal representation - **R**obust **O**rdinal **R**egression - UTA^{GMS}

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

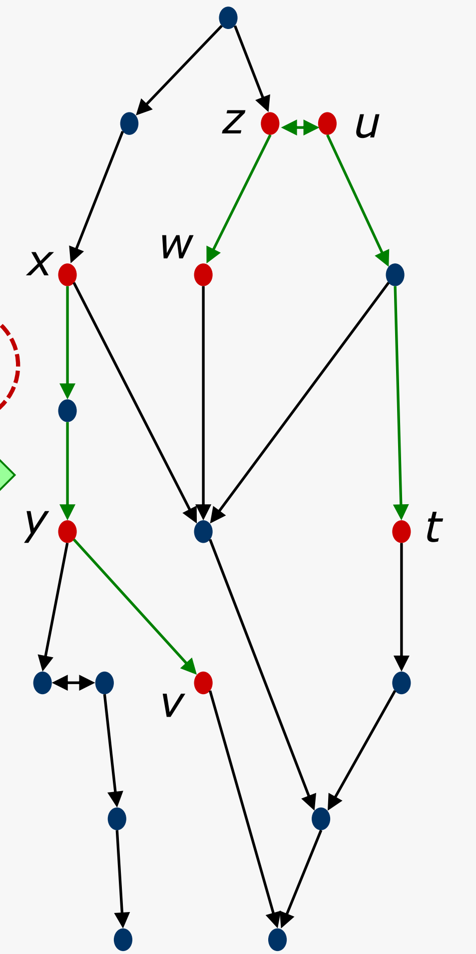
preference information

$x \succsim y$
$z \succsim w$
$y \succsim v$
$u \succsim t$
$z \succsim u$
$u \succsim z$

LP



LP



necessary ranking
(partial preorder)

Non-univocal representation - **R**obust **O**rdinal **R**egression - UTA^{GMS}

additional preference information

$x \succsim y$

$z \succsim w$

$y \succsim v$

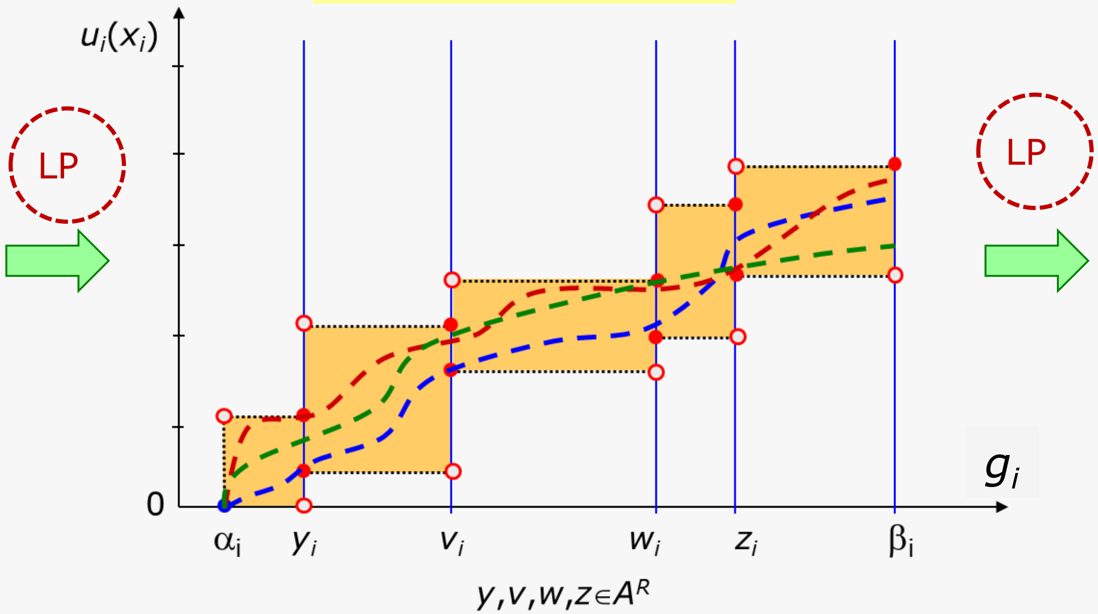
$u \succsim t$

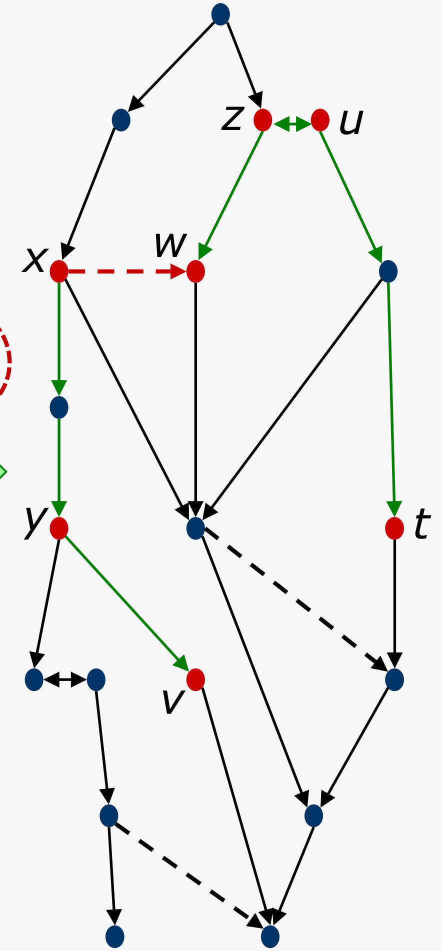
$z \succsim u$

$u \succsim z$

$x \succsim w$

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



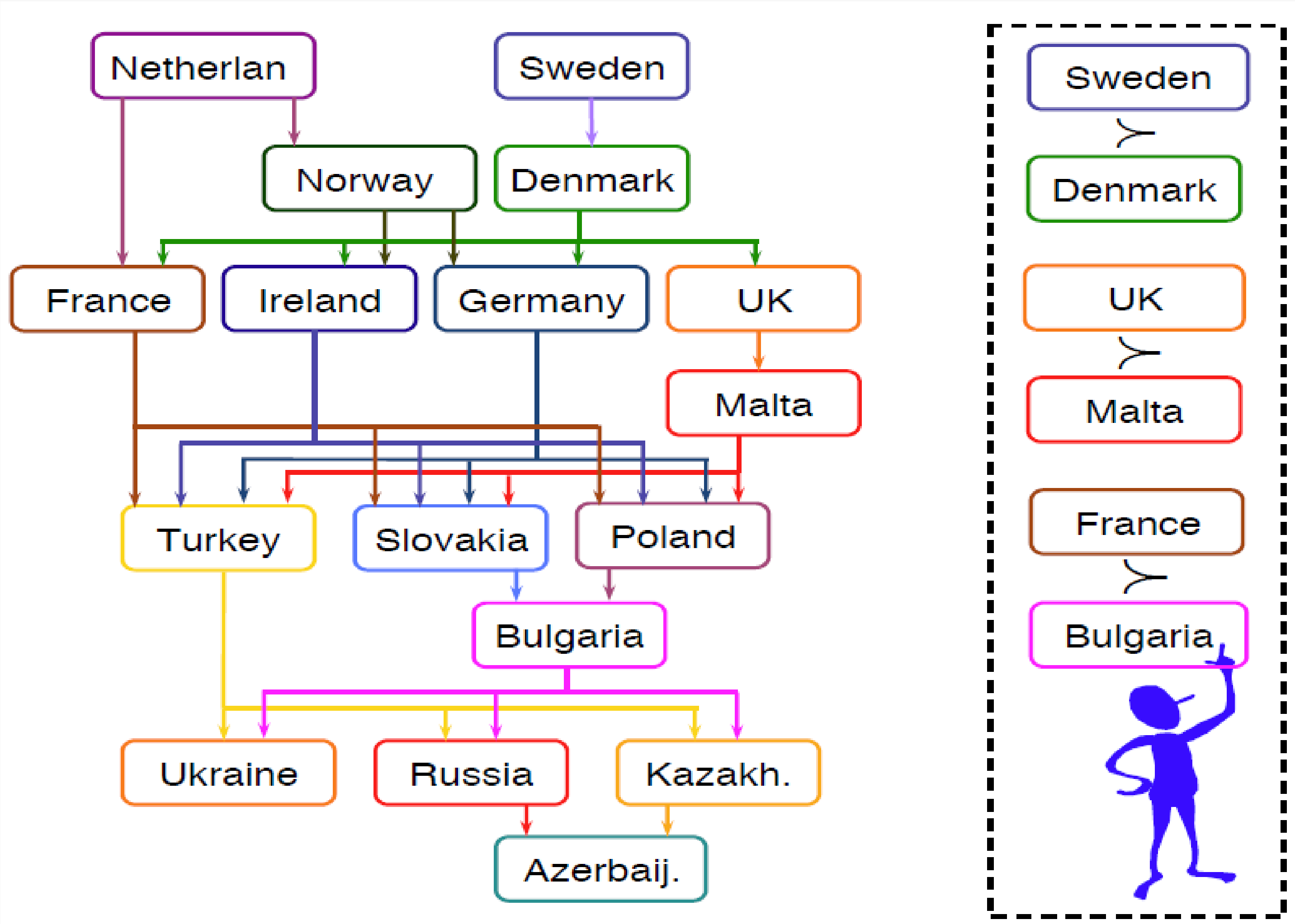


necessary ranking

enriched

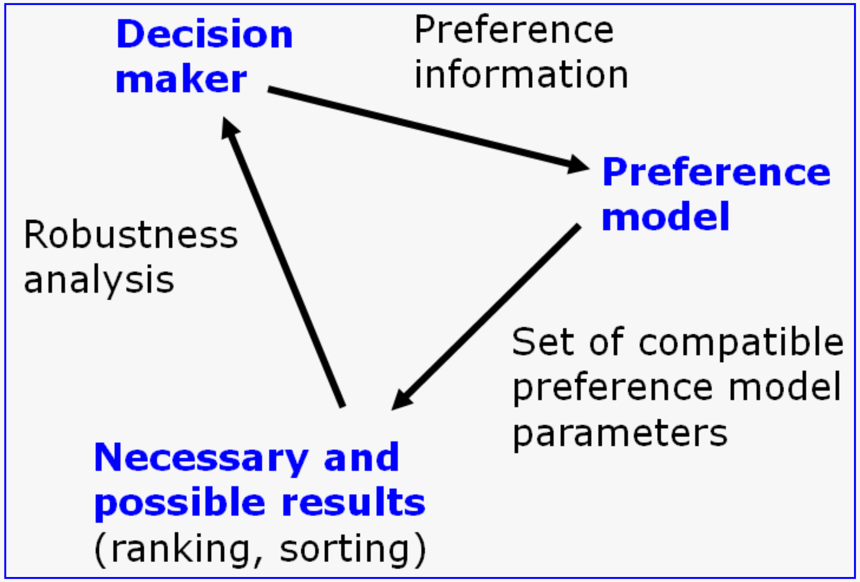
Recommendation in terms of a necessary ranking - UTA^{GMS}

- 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



S. Corrente, S. Greco, M. Kadziński, R. Słowiński: Robust ordinal regression in preference learning and ranking. *Machine Learning*, 93 (2013) 381-422

Checking for the existence of a compatible value function

UTA^{GMS} method

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

$$\left.\begin{array}{l}
\left.\begin{array}{ll}
U(a^*) \geq U(b^*) + \varepsilon & \text{if } a^* \succ b^* \\
U(a^*) = U(b^*) & \text{if } a^* \sim b^*
\end{array}\right\} \forall a^*, b^* \in A^R \\
u_i(x_i^k) - u_i(x_i^{k-1}) \geq 0, \quad i = 1,\ldots,n, \; k = 1,\ldots,m_i(A^R) \\
u_i(x_i^0) = 0, \quad i = 1,\ldots,n \\
\sum_{i=1}^{n} u_i(x_i^{m_i}) = 1
\end{array}\right\} E^{A^R}$$

Since $U(a) = \sum_{i=1}^{n} u_i(a)$, the only unknown of this LP problem are marginal value functions u_i

Checking for the existence of a compatible value function

UTA^{GMS} method

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

$$\left.\begin{array}{l}\left.\begin{array}{l} U(a^*) \geq U(b^*) + \varepsilon \quad \text{if } a^* \succ b^* \\ U(a^*) = U(b^*) \quad \text{if } a^* \sim b^* \end{array}\right\} \forall a^*, b^* \in A^R \\ u_i(x_i^k) - u_i(x_i^{k-1}) \geq 0, \quad i = 1, \ldots, n, \; k = 1, \ldots, m_i(A^R) \\ u_i(x_i^0) = 0, \quad i = 1, \ldots, n \\ \sum_{i=1}^{n} u_i(x_i^{m_i}) = 1 \end{array}\right\} E^{A^R}$$

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 , $i=1,\ldots,n$; then E^{A^R} becomes $E(a,b)$
- Consider constraints:

$$\left.\begin{array}{c} U(b) \geq U(a) + \varepsilon \\ E(a,b) \end{array}\right\} E^N(a,b) \qquad \left.\begin{array}{c} U(a) \geq U(b) \\ E(a,b) \end{array}\right\} E^P(a,b)$$

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

$$a \succeq^N b \Leftrightarrow \text{if } E^N(a,b) \text{ infeasible or } \varepsilon^N(a,b) = \max \varepsilon, \text{ s.t. } E^N(a,b) \text{ is} \leq 0$$

$$a \succeq^P b \Leftrightarrow \text{if } E^P(a,b) \text{ feasible and } \varepsilon^P(a,b) = \max \varepsilon, \text{ s.t. } E^P(a,b) \text{ is} > 0$$

ROR including information about intensities of preference – GRIP

- GRIP extends the UTA^{GMS} method by adopting all features of UTA^{GMS} and by taking into account additional preference information:
  - comprehensive comparisons of intensities of preference between some pairs of reference actions, e.g. „*x* is preferred to *y* at least as much as *w* is preferred to *z*"
  - partial comparisons of intensities of preference between some pairs of reference actions on particular criteria, e.g. „*x* is preferred to *y* at least as much as *w* is preferred to *z*, on criterion $g_i \in F$"

J. Figueira, S. Greco, R. Słowiński: Building a set of additive value functions representing a reference preorder and intensities of preference: GRIP method. *European J. Operational Research*, 195 (2009) 460-486.

Checking for the existence of a compatible value function

GRIP method

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

$$\left.\begin{array}{l}
\left.\begin{array}{l}
U(a^*) \geq U(b^*) + \varepsilon \quad \text{if} \quad a^* \succ b^* \\
U(a^*) = U(b^*) \quad \text{if} \quad a^* \sim b^* \\
U(a^*) - U(b^*) \geq U(c^*) - U(d^*) + \varepsilon \quad \text{if} \quad (a^*, b^*) \succ^* (c^*, d^*) \\
U(a^*) - U(b^*) = U(c^*) - U(d^*) \quad \text{if} \quad (a^*, b^*) \sim^* (c^*, d^*) \\
u_i(a^*) - u_i(b^*) \geq u_i(c^*) - u_i(d^*) + \varepsilon \quad \text{if} \quad (a^*, b^*) \succ_i^* (c^*, d^*), \ i = 1, \ldots, n \\
u_i(a^*) - u_i(b^*) = u_i(c^*) - u_i(d^*) \quad \text{if} \quad (a^*, b^*) \sim_i^* (c^*, d^*), \ i = 1, \ldots, n
\end{array}\right\} \forall a^*, b^*, c^*, d^* \in A^R \\
u_i(x_i^k) - u_i(x_i^{k-1}) \geq 0, \quad i = 1, \ldots, n, \ k = 1, \ldots, m_i(A^R) \\
u_i(x_i^0) = 0, \quad i = 1, \ldots, n; \quad \sum_{i=1}^{n} u_i(x_i^{m_i}) = 1
\end{array}\right\} E_G^{A^R}$$

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

When the adopted value function fails to represent preferences…

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

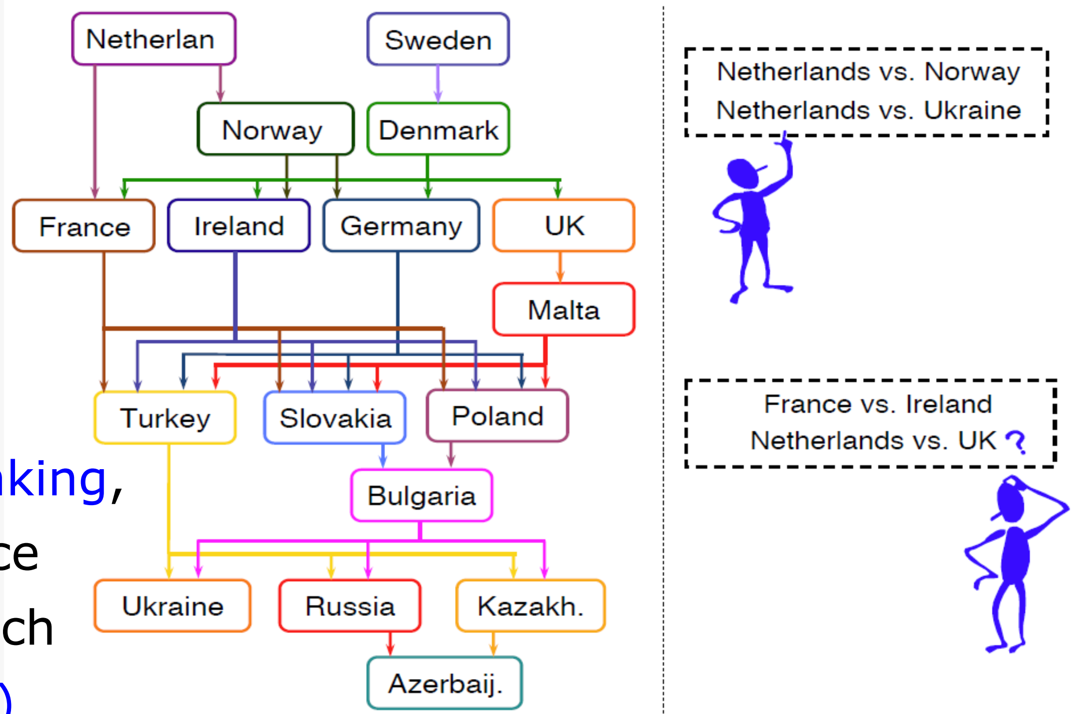
- identify and eliminate „troublesome” pieces of preference information (Mousseau et al. 2003),
- continue to use „not completely compatible” set of value functions with an acceptable approximation error
- augment the complexity of the value function, i.e., pass from additive value function to Choquet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

S. Greco, V. Mousseau, R. Słowiński: UTA^{GMS}–INT: robust ordinal regression of value functions handling interacting criteria. *EJOR*, 239 (2014) 711–730.

Representative instance of the preference model

One can also work with a „representative" value function

- It may be desirable to have a total order and scores of actions
- The idea is to select among compatible value functions that one which better highlights the necessary ranking, i.e., maximizes the difference of values for pairs (a, b), such that $a \succsim^N b$ while not($b \succsim^N a$)

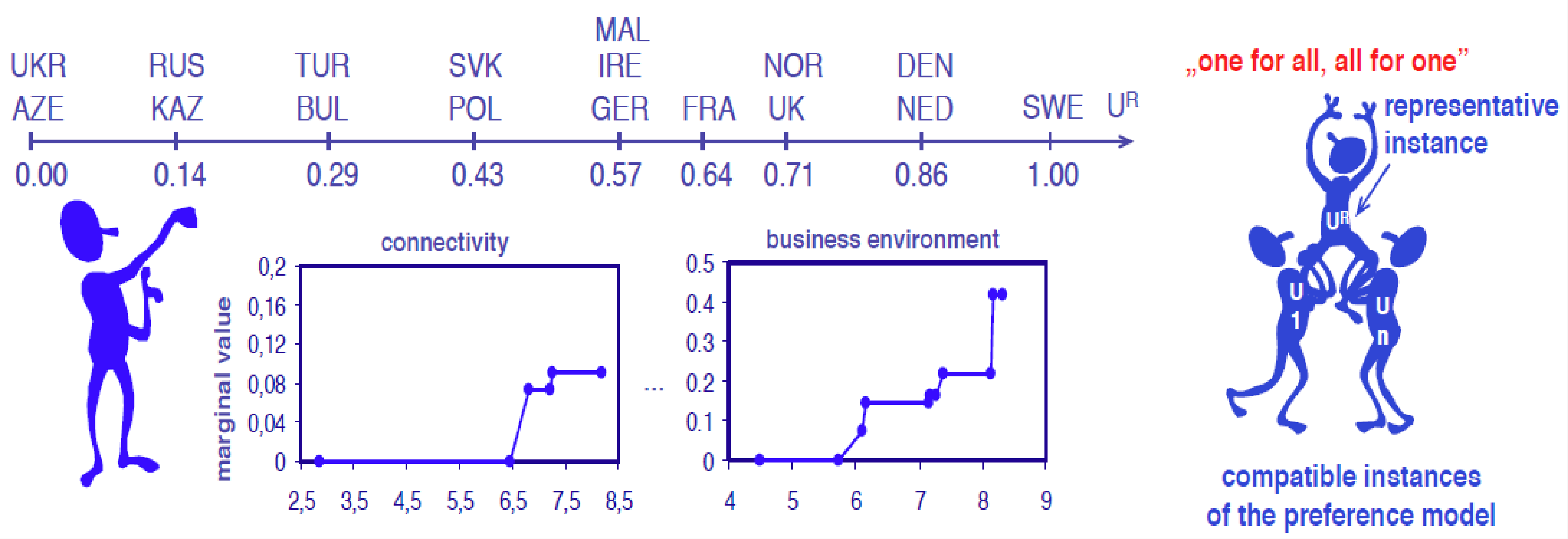


- As secondary objective, we minimize the difference of values for pairs (a, b) for which no necessary relation holds, i.e., such that not($a \succsim^N b$) and not($b \succsim^N a$)

M. Kadziński, S. Greco, R. Słowiński. Selection of a representative value function in robust multiple criteria ranking and choice. *EJOR*, 217 (2012) 541-553

One can also work with a „representative” value function

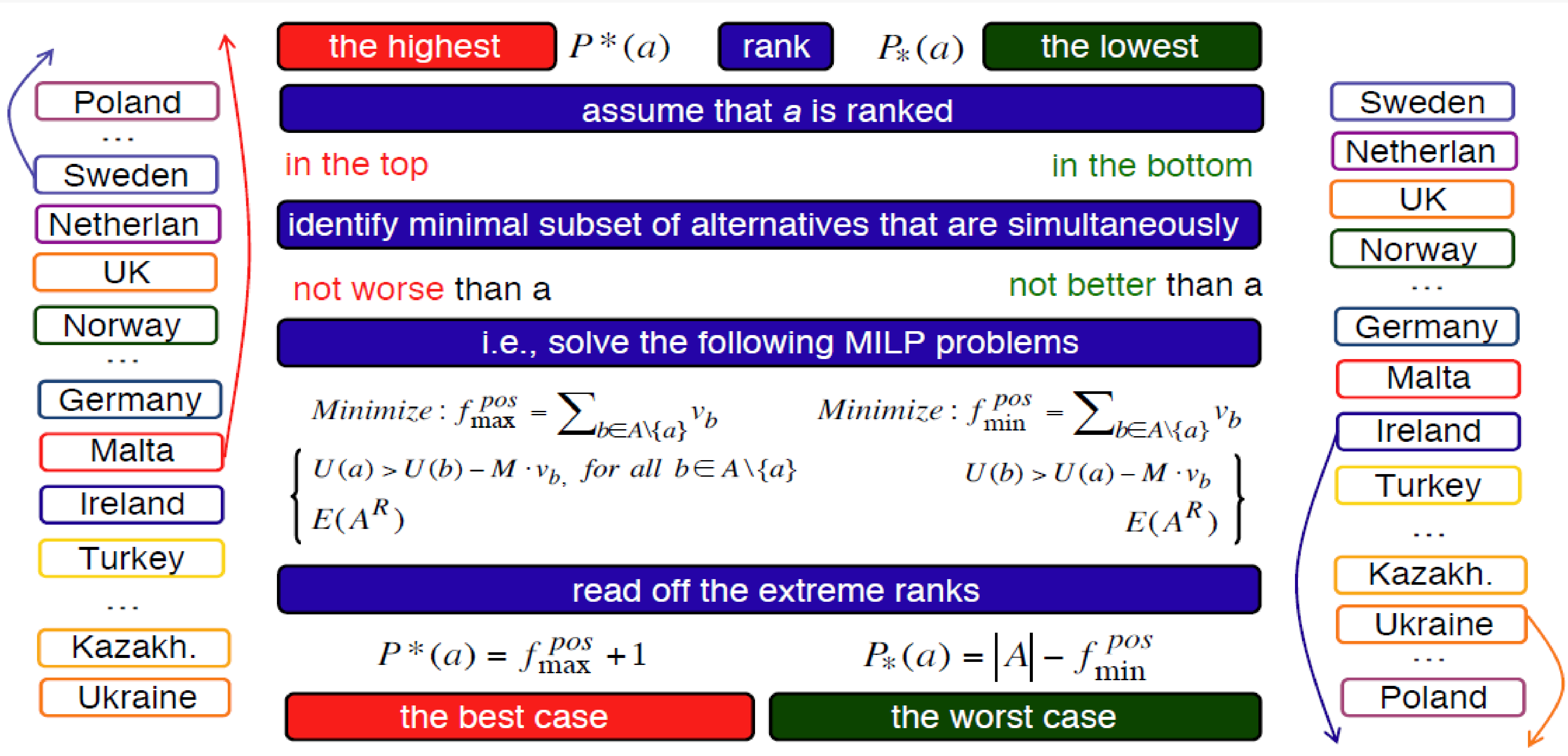
- Reflects a reasonable compromise between all possible outcomes
- Highlights the most stable parts of the ranking



Extreme ranking analysis

Extreme ranking analysis

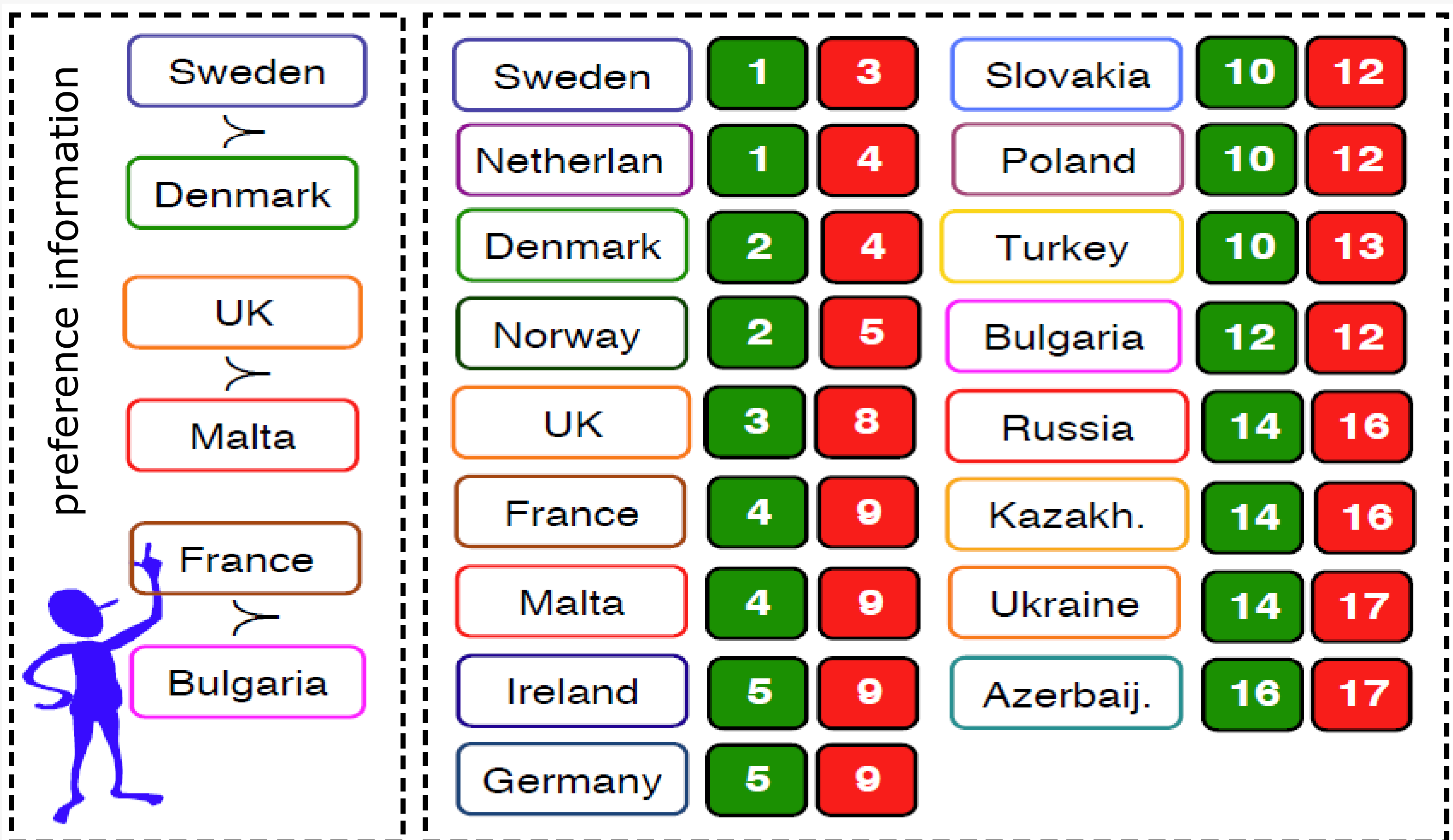
- Collate each action with all the remaining actions jointly
- Compute the highest and the lowest ranks and scores



M. Kadziński, S. Greco, R. Słowiński: Extreme ranking analysis in robust ordinal regression. *OMEGA*, 40 (2012) 488-501

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., BEST = $\{a \in A \colon P^*(a)=1\}$

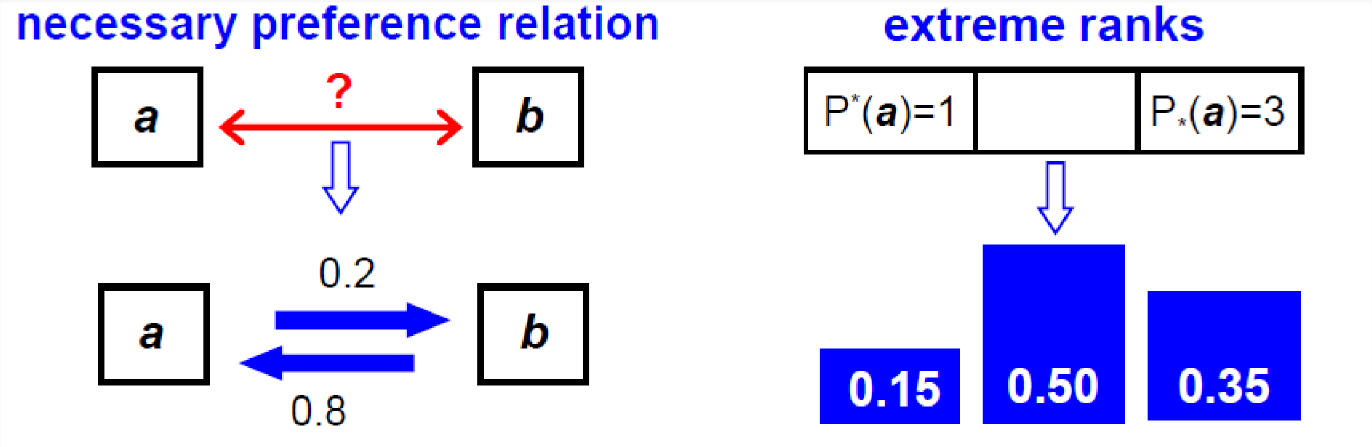


Stochastic ordinal regression

Stochastic Multiobjective Acceptability Analysis & ROR = SOR

- When the necessary preference relation $\succsim^N$ is poor, it leaves many pairs of alternatives incomparable, i.e., $a \not\succsim^P b$ and $b \not\succsim^P a$
- 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)

Stochastic Multiobjective Acceptability Analysis & ROR → SOR

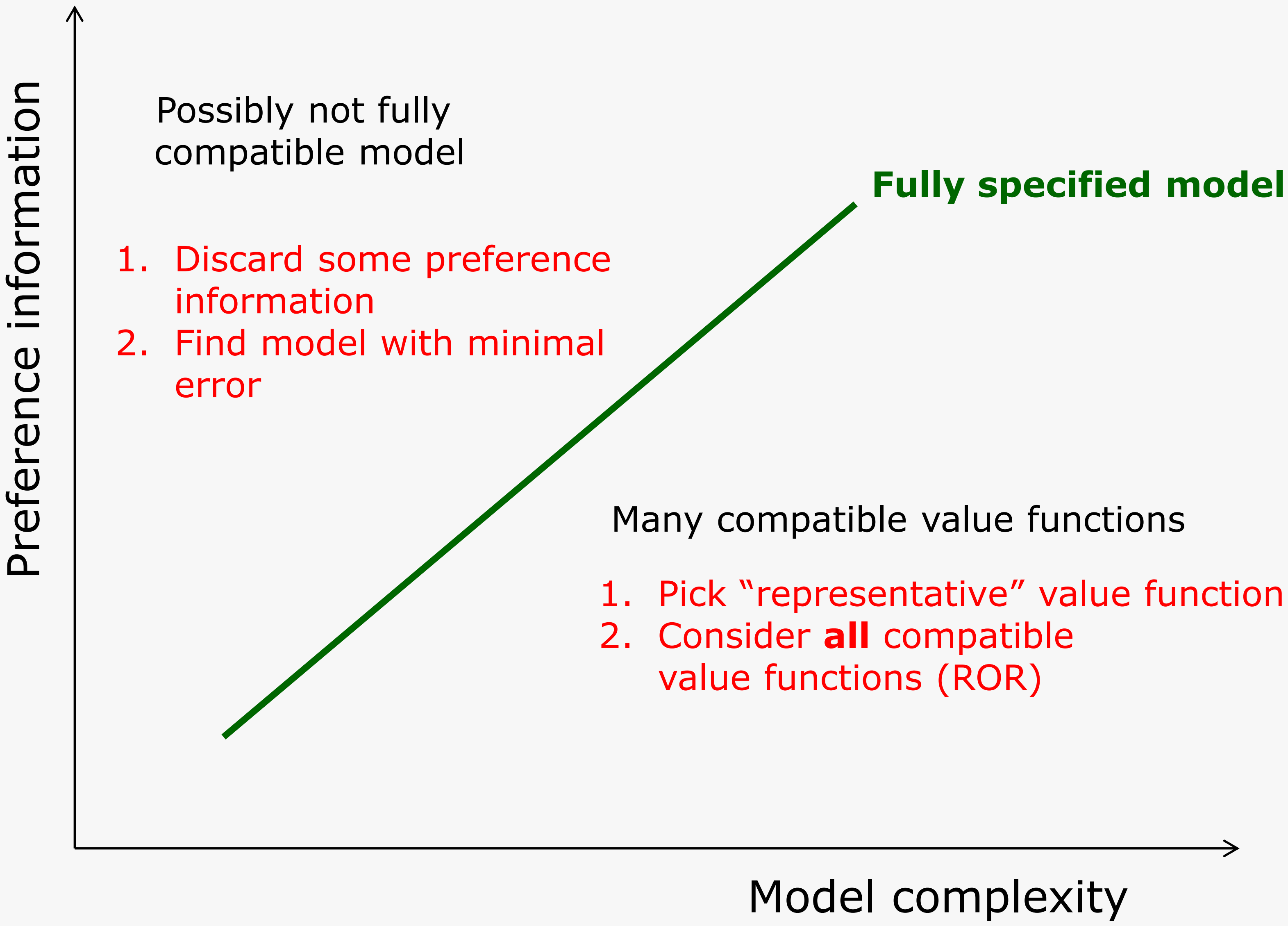


pairwise winning indices   rank acceptability indices

M. Kadziński, T. Tervonen, Stochastic ordinal regression for multiple criteria sorting, *Decision Support Systems*, 55(1), 55-66, 2013

S. Corrente, S. Greco, M. Kadziński, R. Słowiński: Inducing probability distributions on the set of value functions by Subjective Stochastic Ordinal Regression. *Knowledge Based Systems*, 112 (2016) 26–36

Preference information vs. model complexity



Model complexity vs. over-fitting

- Fully specified model is exposed to the risk of over-fitting and may be sensitive to noise
- To ensure a better generalization performance, it is reasonable to learn a preference model in a regularization framework
- Find model U by minimizing the regularized loss function:

$$\min_{U \in \boldsymbol{\mathcal{U}}} \Omega(U) + C \sum_{i=1}^{n} l(U(x_i), y_i)$$

where $\Omega(U)$ is controlling the model complexity (*structural risk*), and $l(U(x_i), y_i)$ is a loss function measuring the deviation between the actual result y_i and the estimated result $U(x_i)$ for any sample (x_i, y_i) (*empirical risk*); C is a trade-off constant

Model complexity vs. over-fitting

- E.g., for additive value function U composed of piecewise-linear marginal value functions (mvf):
  - model complexity $\Omega(U)$ is a "smoothness" of mvf (closeness to linearity),
  - loss function $l(U(x_i), y_i)$ is a value gap $\xi(a, b)$ that satisfies the implication: $a \succ b \Rightarrow U(a) > U(b) - \xi(a, b)$
- A trade-off between model's complexity and its fitting ability is achieved through quadratic optimization
- Non-monotonic criteria (marginal value functions) can be considered

J. Liu, X. Liao, M. Kadziński, R. Słowiński: Preference disaggregation within the regularization framework for sorting problems with multiple potentially non-monotonic criteria. *EJOR*, 276 (2019) 1071–1089

Robust Ordinal Regression
for outranking relation preference model

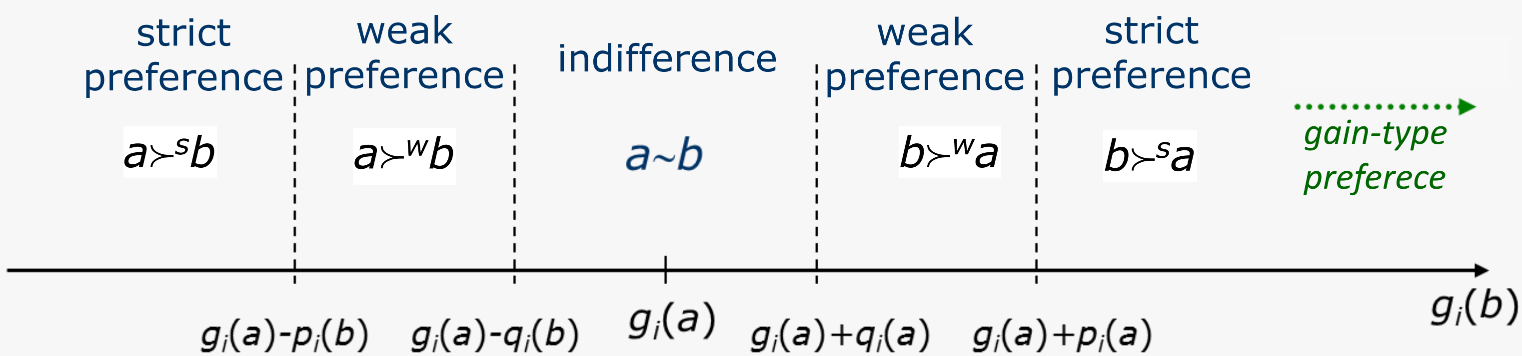
The need for an incomplete and intransitive preference structure

- Value function model is a complete and transitive preference relation with a compensatory logic
- In many real-life decision situations it is reasonable to consider:
  - Incomparability between alternatives (the available information does not permit to compare pairwise all alternatives)
  - Intransitive indifferences (Luce's tea cup paradox) and intransitive preferences (Condorcet paradox)
  - Non-compensatory multicriteria aggregation (what price reduction would you require for a reduction of your car safety by one star?)
- Outranking methods, such as ELECTRE, PROMETHEE, MAPPAC and PRAGMA, answer these needs in Multiple Criteria Decision Aiding

J.R. Figueira, S. Greco, R. Słowiński, B. Roy: An overview of ELECTRE methods and their recent extensions. *Journal of Multi-Criteria Decision Analysis*, 20 (2013) 61–85

Robust Ordinal Regression approach for outranking methods

- Outranking relation S groups three basic preference relations: $\sim$, $\succ^w$, $\succ^s$



aSb reads „alternative a is at least as good as alternative b"

$$aSb \wedge bSa \;\Leftrightarrow\; a \sim b \quad \text{(indifference)}$$

$$aSb \wedge \text{non}(bSa) \;\Leftrightarrow\; a \succ^w b \vee a \succ^s b \quad \text{(large preference)}$$

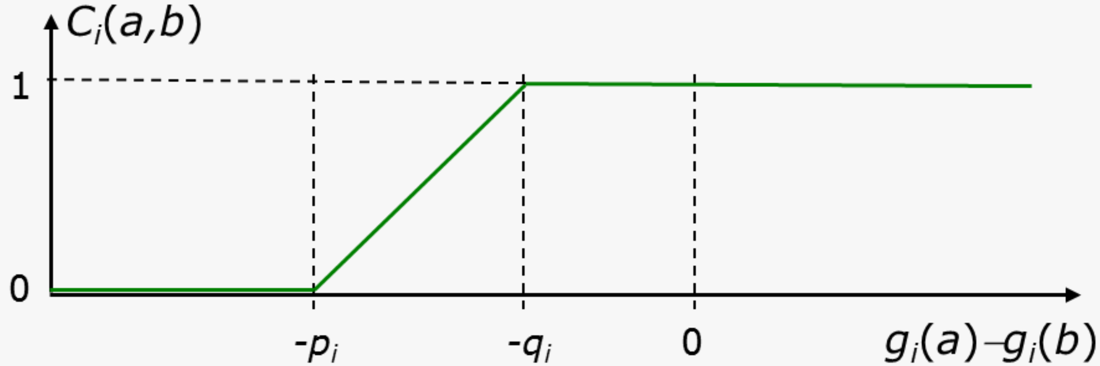
$$\text{non}(aSb) \wedge \text{non}(bSa) \;\Leftrightarrow\; a?b \quad \text{(incomparability)}$$

- S is an incomplete and intransitive relation on set of actions A, constructed via concordance and discordance tests (ELECTRE, Roy 1985)

Robust Ordinal Regression approach for outranking methods

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

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

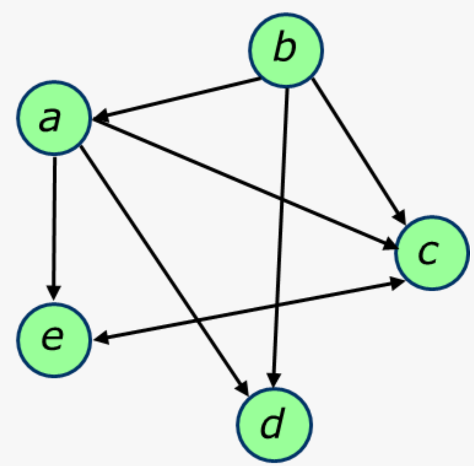


- Concordance test is positive if: $C(a,b)\geq\lambda$, where $\lambda\in[0.5, 1]$ is a cutting level (concordance threshold)
- No compensation between criteria <u>because the weights are not multiplied by performances</u> (weight w_i is a voting power of g_i)

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)

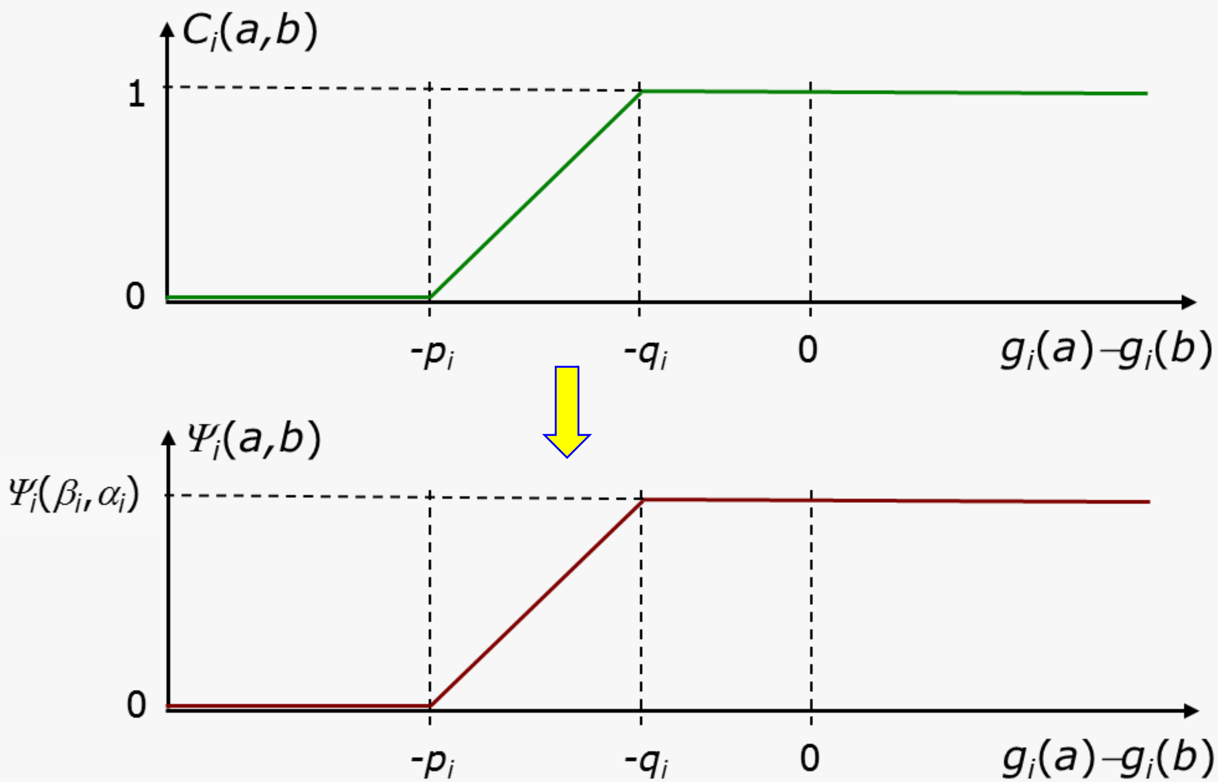
S	*a*	*b*	*c*	*d*	*e*
a	1	0	1	1	1
b	1	1	1	1	0
c	0	0	1	0	1
d	0	0	0	1	0
e	0	0	1	0	1



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 α_i, β_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}):

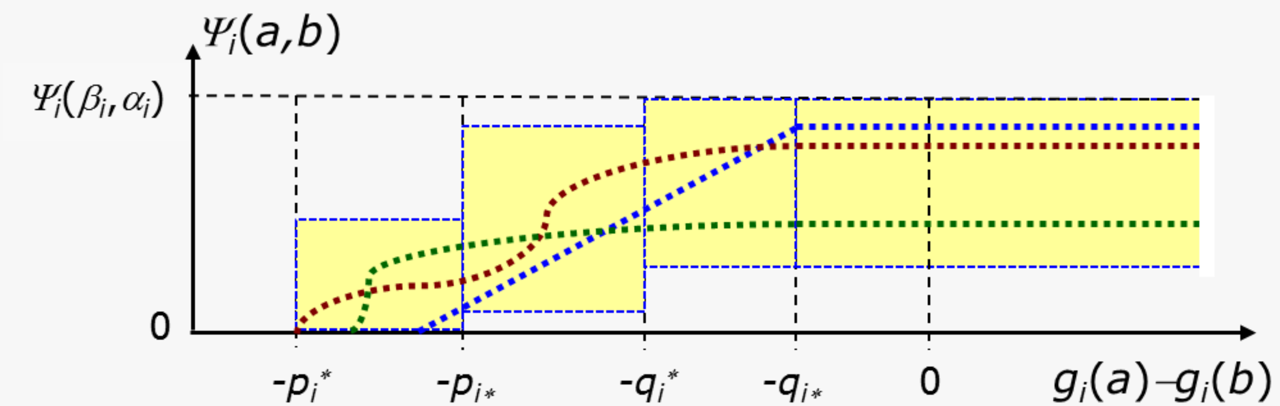
$$aSb \text{ or } aS^{c}b, \text{ for } a,b \in A^{R} \subset A$$

$[q_{i*}, q_i^*]$ - the range of indifference threshold allowed by the DM

$[p_{i*}, p_i^*]$ - the range of preference threshold allowed by the DM

S. Greco, M. Kadziński, V. Mousseau, R. Słowiński: ELECTRE^{GKMS}: Robust ordinal regression for outranking methods. *EJOR*, 214 (2011) 118-135

- Compatible outranking model is a set of marginal concordance functions $\Psi_i(a,b)$, cutting levels λ, indifference q_i, preference p_i, and veto thresholds v_i, $i=1,\ldots,n$, 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^{A^R}:

If aSb for $(a,b)\in A^R\times A^R$:

$$C(a,b)=\sum_{i=1}^{n}\Psi_i(a,b)\geq\lambda$$

$$g_i(b)-g_i(a)+\varepsilon\leq v_i,\ \ i=1,\dots,n$$

aSb
concordance test (+)
and
discordance test (+)

If aS^cb for $(a,b)\in A^R\times A^R$:

$$C(a,b)=\sum_{i=1}^{n}\Psi_i(a,b)+\varepsilon\leq\lambda+M_0(a,b)$$

$$g_i(b)-g_i(a)\geq v_i-\delta M_i(a,b),\ \ i=1,\dots,n$$

$$M_i(a,b)\in\{0,1\},\ \ i=0,1,\dots,n$$

$$\sum_{i=0}^{n}M_i(a,b)\leq n,\ \ \text{where } \delta \text{ is a big given value}$$

aS^cb
concordance test (–)
or
discordance test (–)

$$0.5\leq\lambda\leq1,$$

$$v_i\geq p_i^*+\varepsilon,\ \ \text{if } \left[p_{i*},p_i^*\right] \text{ was given}$$

$v_i\geq g_i(b)-g_i(a)+\varepsilon,\ \ v_i\geq g_i(a)-g_i(b)+\varepsilon,$ if $a\sim_i b$ was given, $i\in\{1,\dots,n\}$

Robust Ordinal Regression approach for outranking methods

- Given a pair of alternatives $a,b \in A$, a possibly outranks b:

$$aS^P b \Leftrightarrow \varepsilon^* > 0$$

where

$$\varepsilon^* = \max \varepsilon$$

$$\text{subject to:}$$

$$\left.\begin{array}{l} E^{A^R} \\ C(a,b) = \sum_{i=1}^{n} \Psi_i(a,b) \geq \lambda \\ g_i(b) - g_i(a) + \varepsilon \leq v_i, \quad i = 1,\ldots,n \end{array}\right\} E^P(a,b)$$

- If $\varepsilon^* > 0$ and constraints $E^P(a,b)$ are feasible, then a outranks b for <u>at least one</u> compatible outranking model ($aS^P b$)

Robust Ordinal Regression approach for outranking methods

- Given a pair of alternatives $a,b \in A$, a necessarily outranks b:

$$aS^N b \;\Leftrightarrow\; \varepsilon^* \le 0$$

where $\varepsilon^* = max\ \varepsilon$

subject to:

$$\left.\begin{array}{l} E^{A^R} \\ C(a,b) = \sum_{i=1}^{n} \Psi_i(a,b) + \varepsilon \le \lambda + M_0(a,b) \\ g_i(b) - g_i(a) \ge v_i - \delta M_i(a,b) \\ M_i(a,b) \in \{0,1\}, \quad i = 1,\dots,n, \quad \sum_{i=0}^{n} M_i(a,b) \le n \end{array}\right\} E^N(a,b)$$

- If $\varepsilon^* \le 0$ or constraints $E^N(a,b)$ are infeasible,

  then a outranks b for <u>all</u> compatible outranking models

  ($aS^N b$ because $aS^{CN} b$ is not possible)

Robust Ordinal Regression approach for outranking methods

- For any pair of alternatives $(a,b) \in A \times A$:

$$aS^Nb \;\Leftrightarrow\; \text{not}(aS^{CP}b), \text{ as well as, } aS^{CP}b \;\Leftrightarrow\; \text{not}(aS^Nb),$$

$$aS^Pb \;\Leftrightarrow\; \text{not}(aS^{CN}b), \text{ as well as, } aS^{CN}b \;\Leftrightarrow\; \text{not}(aS^Pb)$$

so, only aS^Nb and aS^Pb are to be checked

- Thus, there are 2 „sources of information” about 4 relations in A:

$$\boxed{S^N, S^{CN}, S^P, S^{CP}}$$

- Some properties:

$$aS^Nb \;\Rightarrow\; aS^Pb$$

$$aS^Nb \;\Rightarrow\; \text{not}(aS^{CN}b), \text{ as well as, } aS^{CN}b \;\Rightarrow\; \text{not}(aS^Nb)$$

$$aSb \Rightarrow aS^Nb$$

$$aSb \Rightarrow \text{not}(bS^Pa)$$

Exploitation of outranking relations S^N, S^{CN}, S^P, S^{CP} in set A

- Choice problem:

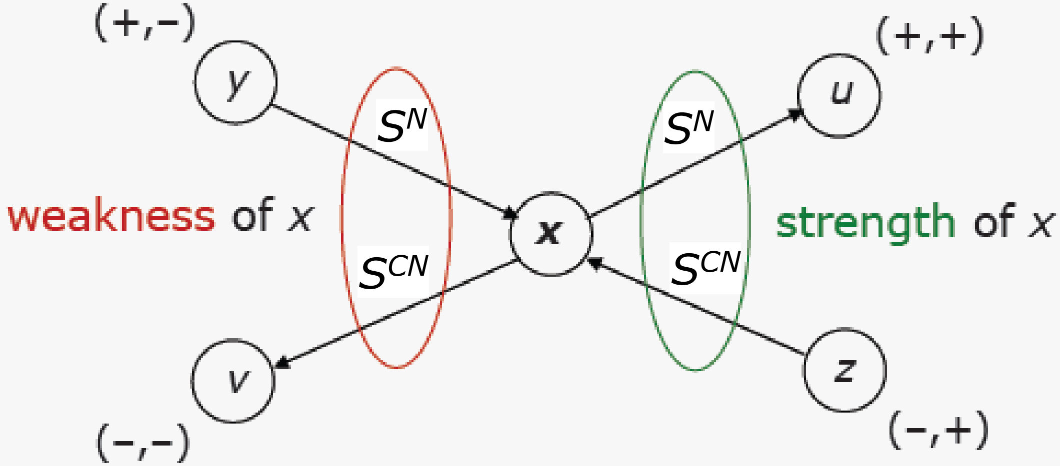
  Kernel of the necessary outranking graph S^N

- 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

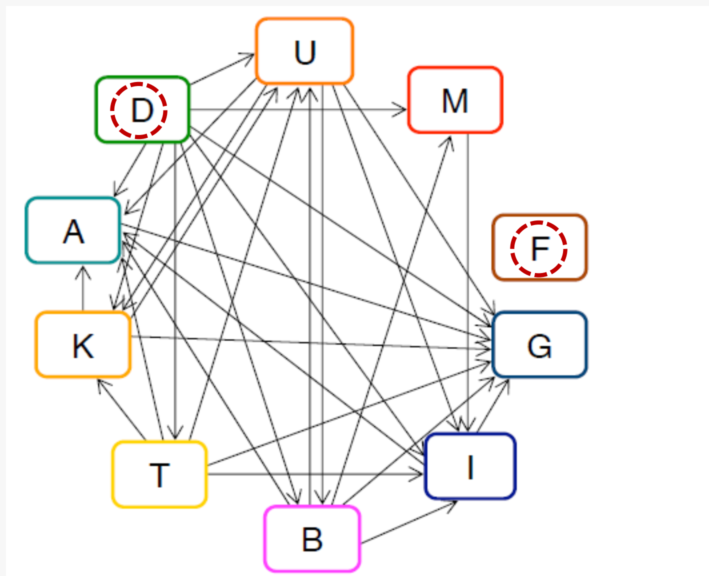
S^N	D	U	M	F	G	I	B	T	K	A
D	1	1	1	0	1	0	1	1	1	1
U	0	1	0	0	1	0	1	0	1	1
M	0	0	1	0	0	0	0	0	0	0
F	0	0	0	1	0	0	0	0	0	0
G	0	0	0	0	1	0	0	0	0	0
I	0	0	0	0	1	1	0	0	0	1
B	0	1	0	0	1	1	1	0	0	1
T	0	1	0	0	1	1	0	1	1	1
K	0	0	0	0	0	0	0	0	1	0
A	0	0	0	0	1	0	0	0	0	1

I S A
T S U
A S^C D
G S^C I

S^N	D	U	M	F	G	I	B	T	K	A
D	1	1	1	0	1	1	1	1	1	1
U	0	1	0	0	1	1	1	0	1	1
M	0	0	1	0	0	1	0	0	0	0
F	0	0	0	1	0	0	0	0	0	0
G	0	0	0	0	1	0	0	0	0	0
I	0	0	0	0	1	1	0	0	0	1
B	0	1	1	0	1	1	1	0	0	1
T	0	1	0	0	1	1	0	1	1	1
K	0	1	0	0	1	0	0	0	1	1
A	0	0	0	0	1	0	0	0	0	1

I S A
T S U
A S^C D
G S^C I
B S M
M S I
U S^C T

kernel



NFS ranking

D
T
F
B
U
K
I
A
M
G

S. Greco, M. Kadziński, V. Mousseau, R. Słowiński:
ELECTRE^GKMS: Robust ordinal regression for outranking methods. *EJOR*, 214 (2011) 118-135

Robust Ordinal Regression approach for outranking methods

- Other developments in ROR for outranking methods in ranking:
  - PROMETHEE^{GKS} and extreme ranking analysis
  - Representative instance of a compatible outranking relations
  - Multiple Criteria Hierachy Process (MCHP) for outranking methods
  - MCHP for ELECTRE III with interacting criteria and Stochastic ROR

Robust Ordinal Regression
for decision rule preference model

Syntax of monotonic decision rules

ordinal classifi-cation

- $if\ \ x_{q1}\succeq_{q1}r_{q1}\ and\ x_{q2}\succeq_{q2}r_{q2}\ and\ \ldots\ x_{qp}\succeq_{qp}r_{qp},\ then\ x\rightarrow$ class t or better
- $if\ \ x_{q1}\preceq_{q1}r_{q1}\ and\ x_{q2}\preceq_{q2}r_{q2}\ and\ \ldots\ x_{qp}\preceq_{qp}r_{qp},\ then\ x\rightarrow$ class t or worse

choice ranking cardinal criteria

- $if\ \ (x\succ_{q1}^{\geq h(q1)}y)\ and\ (x\succ_{q2}^{\geq h(q2)}y)\ and\ \ldots\ (x\succ_{qp}^{\geq h(qp)}y),\ then\ xSy$
- $if\ (x\succ_{q1}^{\leq h(q1)}y)\ and\ (x\succ_{q2}^{\leq h(q2)}y)\ and\ \ldots\ (x\succ_{qp}^{\leq h(qp)}y),\ then\ xS^{c}y$

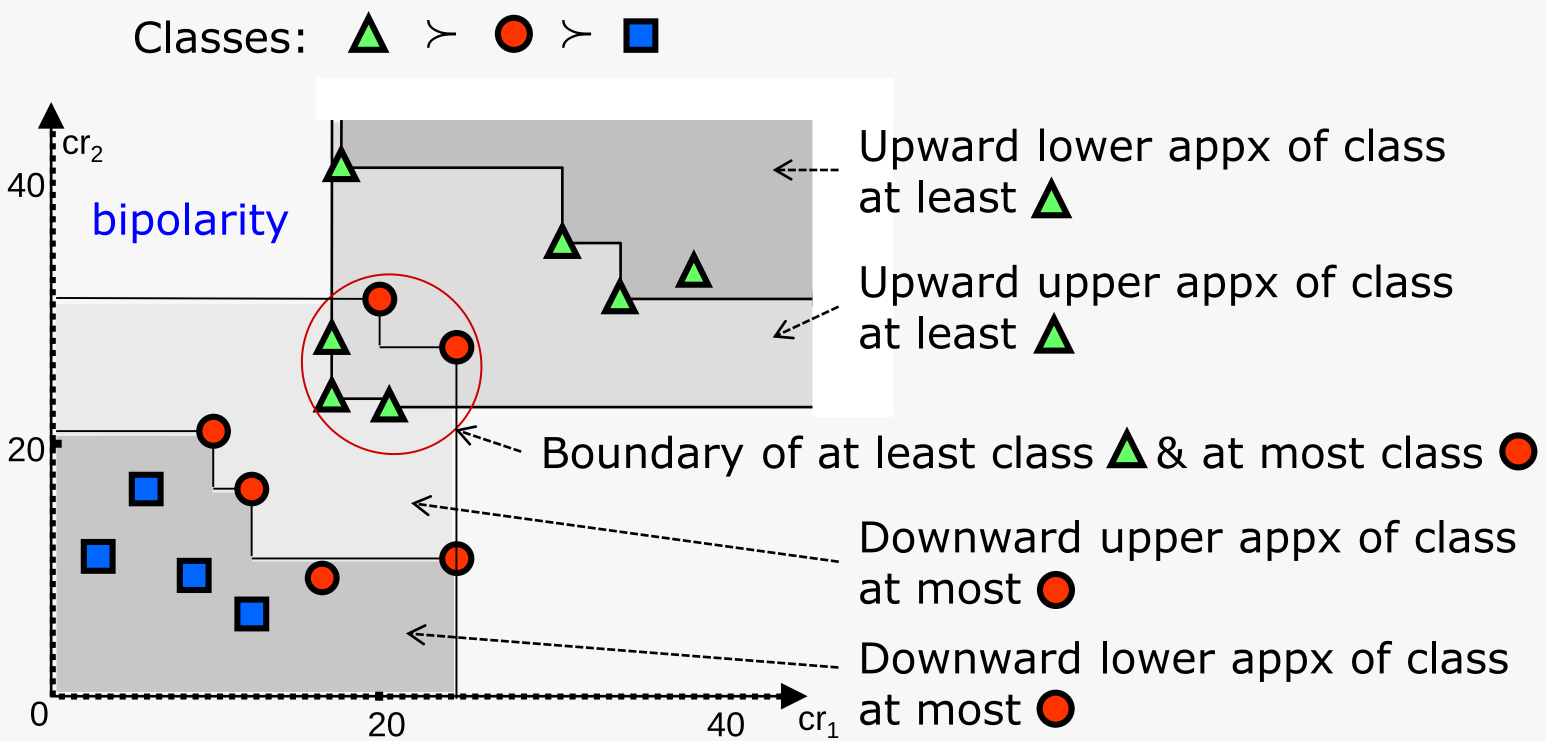
choice ranking ordinal criteria

- $if\ \boxed{x_{g1}\succeq_{g1}r_{q1}\ \&\ y_{g1}\preceq_{g1}r'_{q1}}\ \&\ \ldots\ x_{gp}\succeq_{gp}r_{gp}\ \&\ y_{gp}\preceq_{gp}r'_{gp},\ \ then\ xSy$
- $if\ \boxed{x_{g1}\preceq_{g1}r_{q1}\ \&\ y_{g1}\succeq_{g1}r'_{q1}}\ \&\ \ldots\ x_{gp}\preceq_{gp}r_{gp}\ \&\ y_{gp}\succeq_{gp}r'_{gp},\ \ then\ xS^{c}y$

pair of objects x,y evaluated on criterion g_1

S.Greco, B.Matarazzo, R.Słowiński: Decision rule approach. Chapter 13 [in]: S.Greco M.Ehrgott, J.Figueira (eds.), *Multiple Criteria Decision Analysis: State of the Art Surveys*, 2[nd] edition, OR & MS 233, Springer, New York, 2016, pp. 497-552

Dominance-based Rough Set Approach (DRSA)



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*

S.Greco, B.Matarazzo, R.Słowiński: Rough sets theory for multicriteria decision analysis. *European J. of Operational Research*, 129 (2001) no.1, 1-47

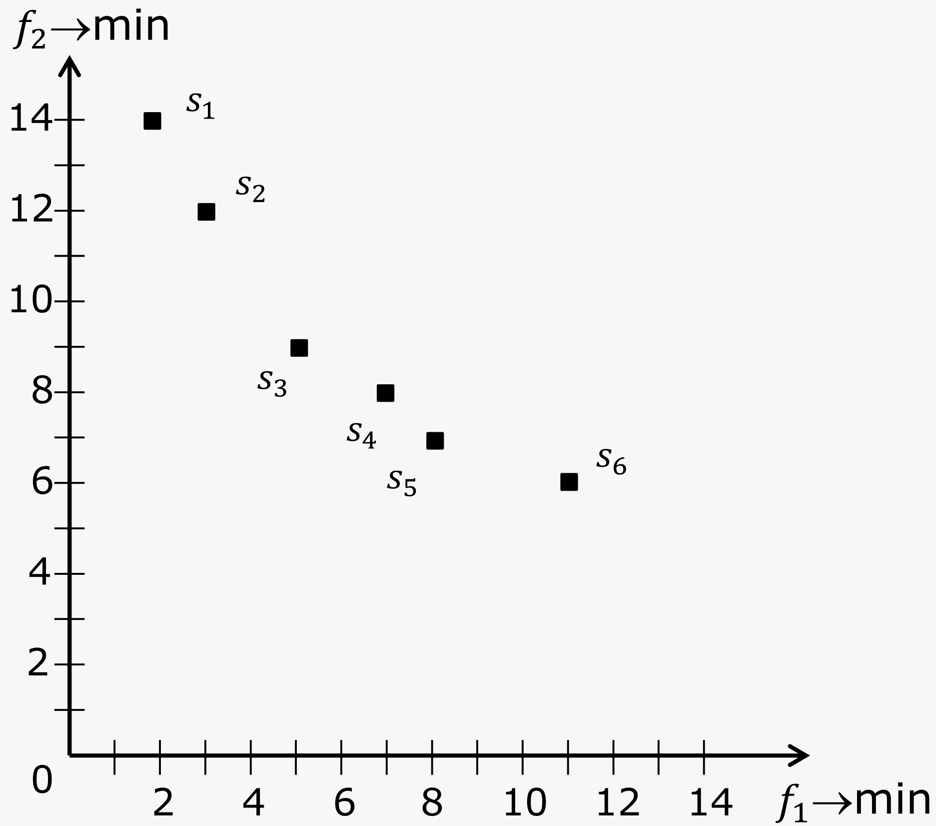
Preference modelling by dominance-based decision rules

- 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 evalautions 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

Example

Sample of 6 n-d solutions submitted to evaluation of the DM

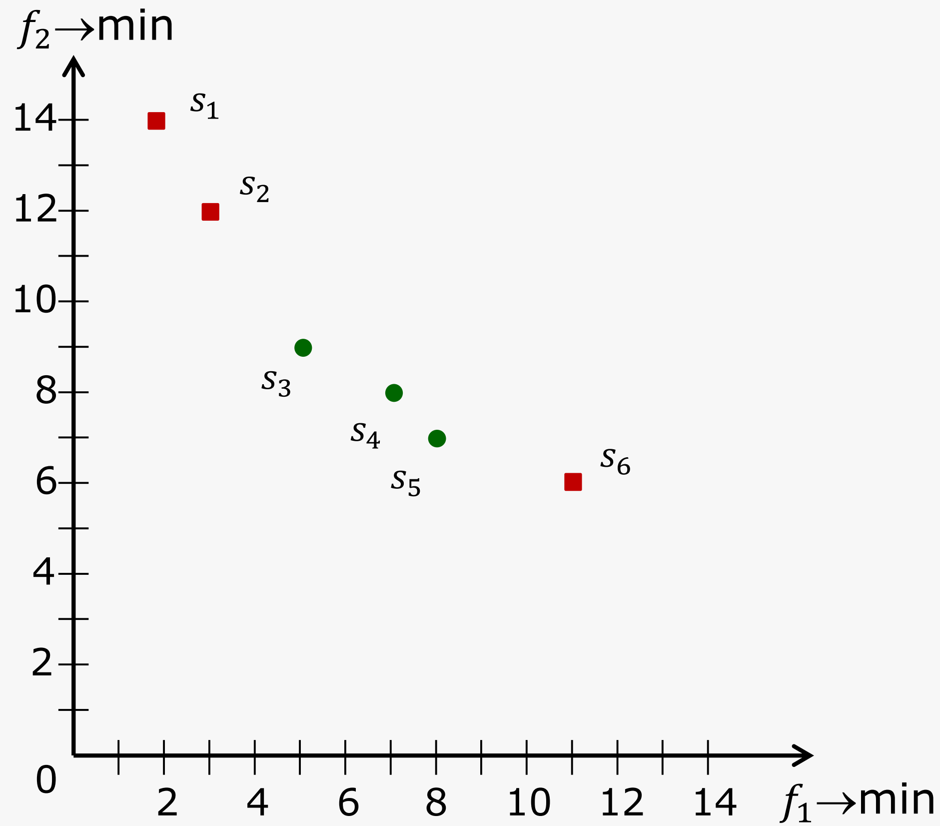
Reference actions	f_1	f_2	DM
s_1	2	14	
s_2	3	12	
s_3	5	9	
s_4	7	8	
s_5	8	7	
s_6	11	6	



Example

Sample of 6 n-d solutions – elicitation of preferences by the DM

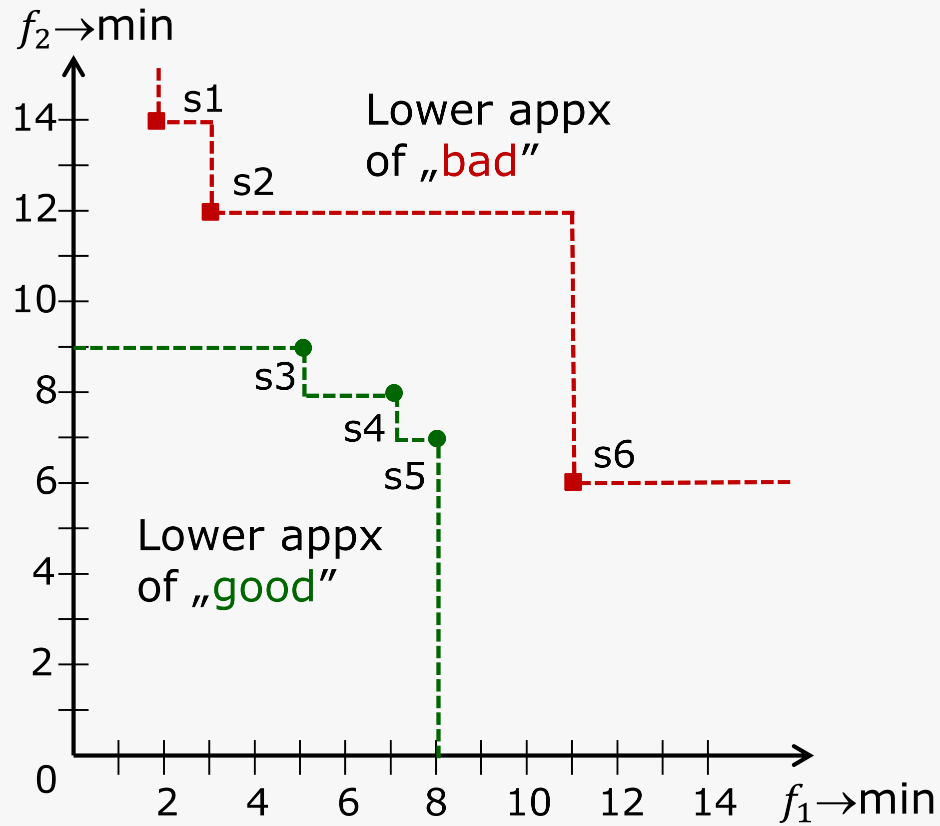
Reference actions	f_1	f_2	DM
s_1	2	14	bad
s_2	3	12	bad
s_3	5	9	good
s_4	7	8	good
s_5	8	7	good
s_6	11	6	bad



Example

Sample of 6 n-d solutions – dominance-based lower approximations

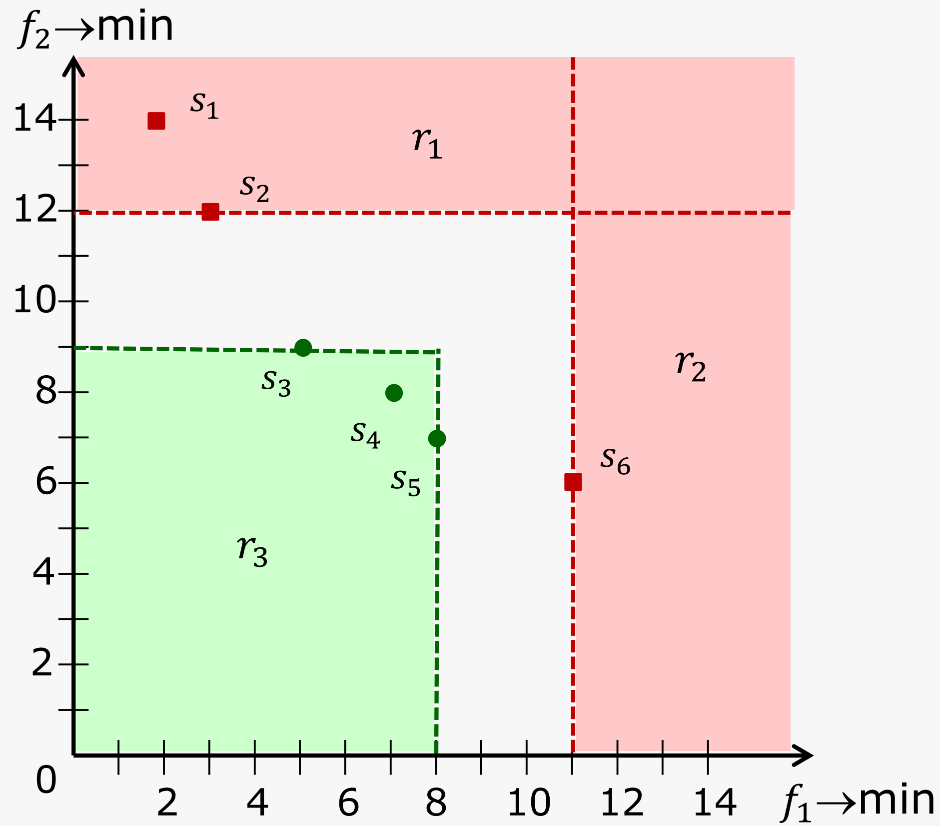
Reference actions	f_1	f_2	DM
s_1	2	14	bad
s_2	3	12	bad
s_3	5	9	good
s_4	7	8	good
s_5	8	7	good
s_6	11	6	bad



Example

Sample of 6 n-d solutions – induction of minimal decision rules

Reference actions	f_1	f_2	DM
s_1	2	14	bad
s_2	3	12	bad
s_3	5	9	good
s_4	7	8	good
s_5	8	7	good
s_6	11	6	bad



$D_{\leq}$ r_1: if $f_2(s) \geq 12$, then s is bad — supported by $\{s_1, s_2\}$

$D_{\leq}$ r_2: if $f_1(s) \geq 11$, then s is bad — supported by $\{s_6\}$

$D_{\geq}$ r_3: if $f_1(s) \leq 8$ & $f_2(s) \leq 9$, then s is good — supported by $\{s_3, s_4, s_5\}$

Examples of applications

Example – Prime d'Excellence Scientifique (PES) *with jMAF*

- Multiple criteria classification of candidates for PES award:
  1. Comprehensive assessment (Global)
  2. Publications (Avis 1)
  3. Supervision of PhD students (Avis 2)
  4. International impact (Avis 3)
  5. Administrative responsibility (Avis 4)

Example – Prime d'Excellence Scientifique (PES) *with jMAF*

Attributes: 5 Examples: 118

No	Global (+)	Avis_1 (+)	Avis_2 (+)	Avis_3 (+)	Avis_4 (+)	PRIME (+)
37	B	A	B	C	B	0
38	B	A	B	B	B	1
39	B	A	A	B	B	1
40	B	A	B	C	B	0
41	B	A	B	B	B	1
42	B	B	B	B	B	0
43	B	A	B	C	B	0
44	B	A	B	B	C	0
45	B	B	B	B	C	0
46	B	A	A	C	B	1
47	B	B	A	B	B	0
48	B	A	A	B	A	1
49	B	B	B	B	B	0
50	B	B	C	C	C	0
51	B	A	B	A	B	1
52	B	A	C	B	C	0

Example – Prime d’Excellence Scientifique (PES) *with jMAF*

Quality of approximation: 0.975

Union name	Accuracy	Cardina...
At most 0	0.962	79
Lower		77
Upper		80
Boundary		3
Example_23		
Example_31		
Example_47		
At least 1	0.927	39
Lower		38
Upper		41
Boundary		3

Name	Cardinality	Content
Core	4	Avis_1, Avis_2, Avis_3, Avis_4
Reducts	1	
Reduct 1	4	Avis_1, Avis_2, Avis_3, Avis_4

Example – Prime d'Excellence Scientifique (PES) *with jMAF*

Number of rules: 7

certain rules

ID	DECISION PART 1	<=	CONDITION 1		CONDITION 2		CONDITION 3
1	(PRIME >= 1)	<=	(Avis_2 >= B)	&	(Avis_4 >= A)		
2	(PRIME >= 1)	<=	(Avis_1 >= A)	&	(Avis_2 >= A)		
3	(PRIME >= 1)	<=	(Avis_1 >= A)	&	(Avis_3 >= B)	&	(Avis_4 >= B)
4	(PRIME <= 0)	<=	(Global <= B)	&	(Avis_4 <= C)		
5	(PRIME <= 0)	<=	(Avis_1 <= B)	&	(Avis_3 <= C)		
6	(PRIME <= 0)	<=	(Avis_2 <= B)	&	(Avis_3 <= C)	&	(Avis_4 <= B)
7	(PRIME <= 0)	<=	(Avis_1 <= B)	&	(Avis_2 <= B)	&	(Avis_4 <= B)

Console | Reducts of PES_RS_var5.isf | Monotonic Unions | Statistics of PES_RS_var5.rules

Rule type: **CERTAIN** Usage type: **AT LEAST** Characteristic class: 1

Support:	28
SupportingExamples:	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 26, 30, 38, 39, 41, 48, 51, 65
Strength:	0.237
Confidence:	1
CoverageFactor:	0.718
Coverage:	28
CoveredExamples:	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 26, 30, 38, 39, 41, 48, 51, 65
NegativeCoverage:	0
InconsistencyMeasure:	0
f-ConfirmationMeasure:	1
A-ConfirmationMeasure:	0.63
Z-ConfirmationMeasure:	1

Example – Prime d'Excellence Scientifique (PES) *with jMAF*

possible rules

ID	DECISION PART 1	<=	CONDITION 1		CONDITION 2		CONDITION 3
8	(PRIME >= 1)	<=	(Avis_2 >= B)	&	(Avis_4 >= A)		
9	(PRIME >= 1)	<=	(Avis_2 >= A)	&	(Avis_3 >= B)		
10	(PRIME >= 1)	<=	(Avis_1 >= A)	&	(Avis_2 >= A)		
11	(PRIME >= 1)	<=	(Avis_1 >= A)	&	(Avis_3 >= B)	&	(Avis_4 >= B)
12	(PRIME <= 0)	<=	(Global <= B)	&	(Avis_4 <= C)		
13	(PRIME <= 0)	<=	(Global <= C)	&	(Avis_3 <= C)		
14	(PRIME <= 0)	<=	(Avis_1 <= B)	&	(Avis_4 <= B)		
15	(PRIME <= 0)	<=	(Avis_2 <= B)	&	(Avis_3 <= C)	&	(Avis_4 <= B)

Console | Reducts of PES_RS_var5.isf | Monotonic Unions | Statistics of PES_RS_var5.rules

Rule type: **POSSIBLE** Usage type: **AT LEAST** Characteristic class: 1

Support: 24
SupportingExamples: 1, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 23, 28, 30, 33, 39, 48, 81
Strength: 0.203
Confidence: 0.923
CoverageFactor: 0.615
Coverage: 26
CoveredExamples: 1, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 23, 28, 30, 31, 33, 39, 47, 48, 81
NegativeCoverage: 2
NegativeCoveredExamples: 31, 47
InconsistencyMeasure: 0.025
f-ConfirmationMeasure: 0.921
A-ConfirmationMeasure: 0.507
Z-ConfirmationMeasure: 0.885

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

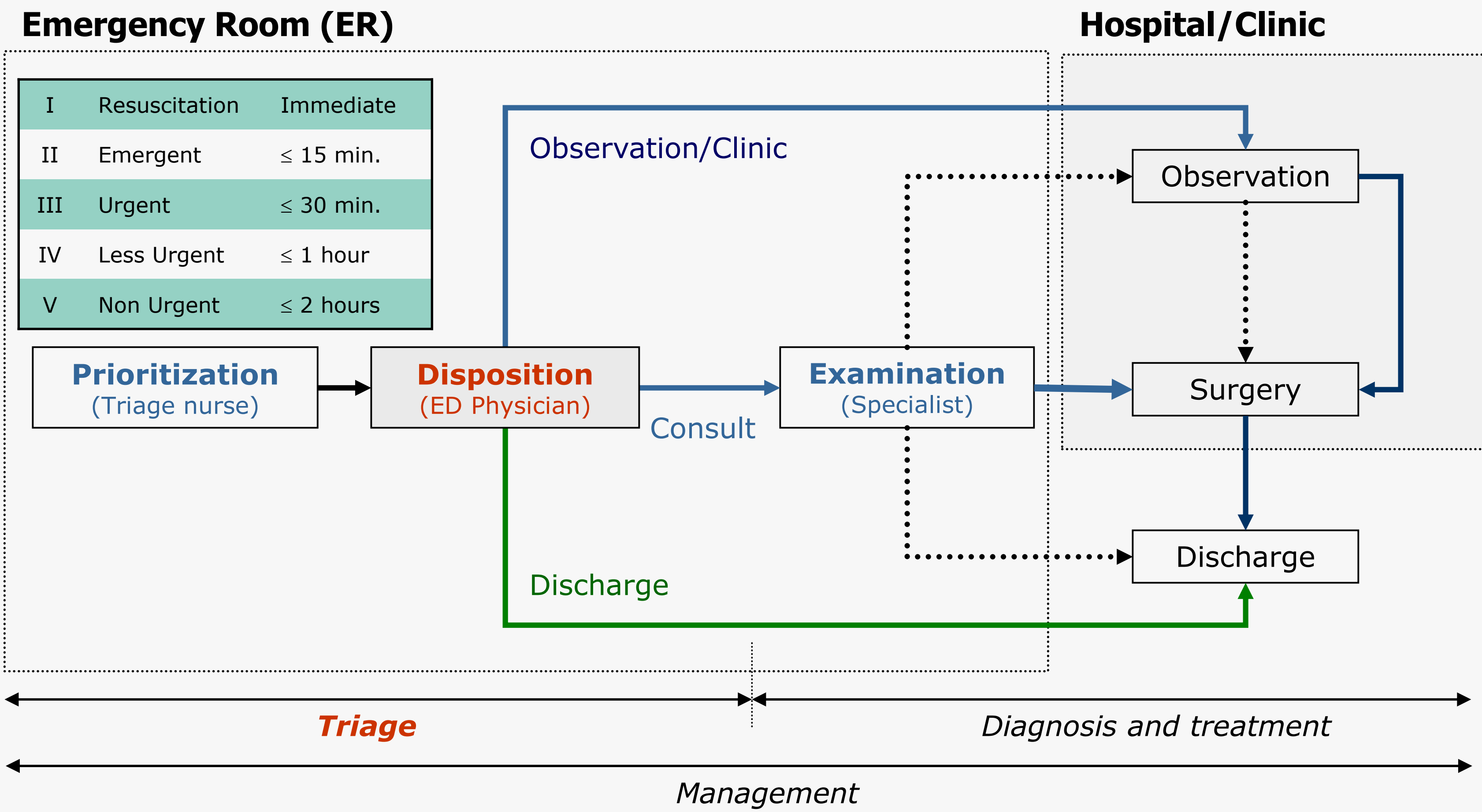
Trial Location

Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

- 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

I	Resuscitation	Immediate
II	Emergent	≤ 15 min.
III	Urgent	≤ 30 min.
IV	Less Urgent	≤ 1 hour
V	Non Urgent	≤ 2 hours



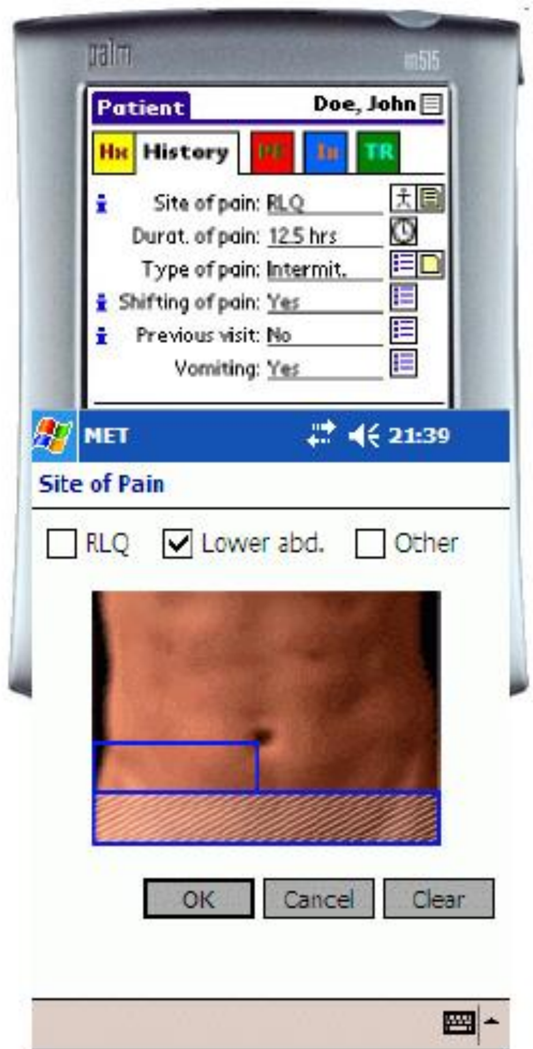
Decision Rules (examples)

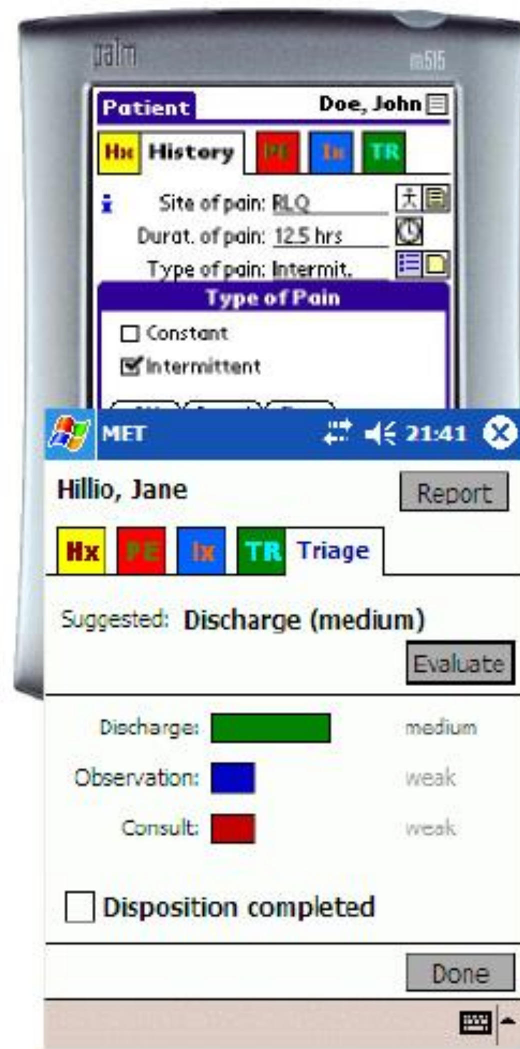
- **if** (Age < 5 years) **and** (PainSite = lower_abdomen)
  and (RebTend = yes) **and** (4 < WBC < 12)
  then (Triage = discharge)

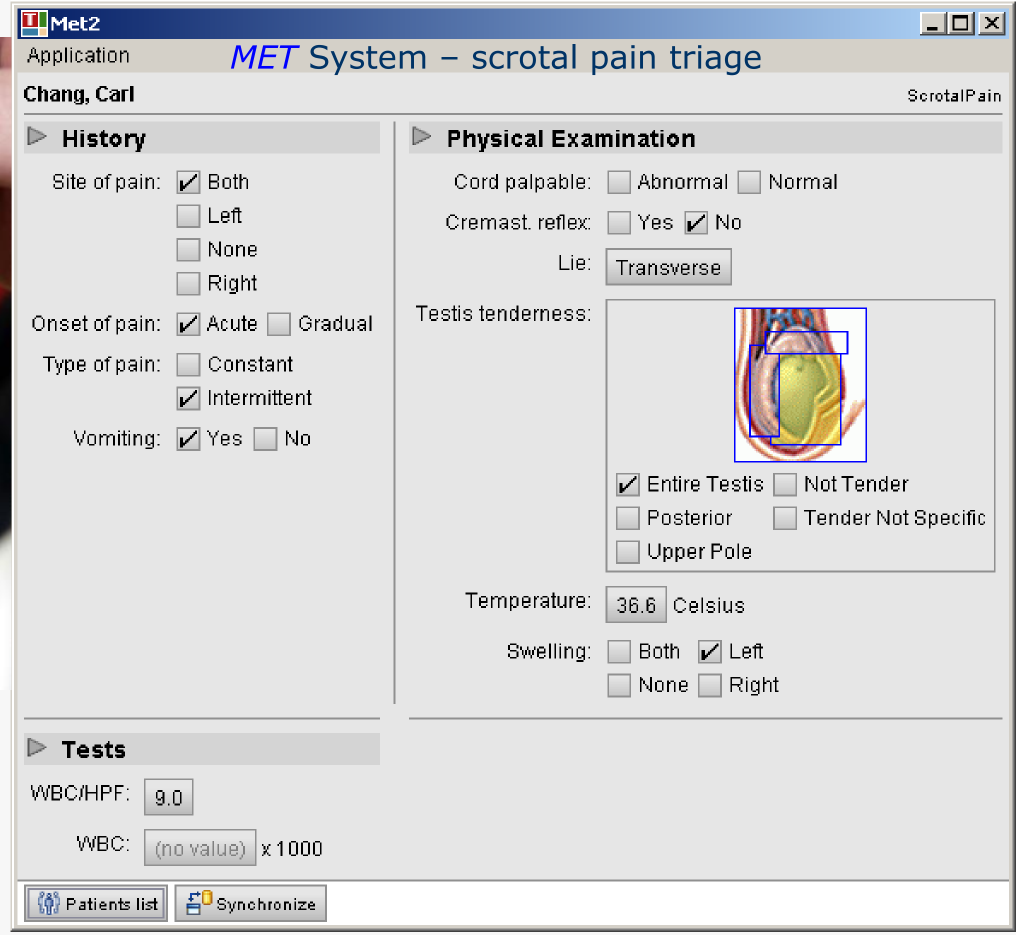
- **if** (PainDur > 7 days) **and** (PainSite = lower_abdomen)
  and (37 ≤ Tempr ≤ 39) **and** (TendSite = lower_abdomen)
  then (Triage = observation)

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

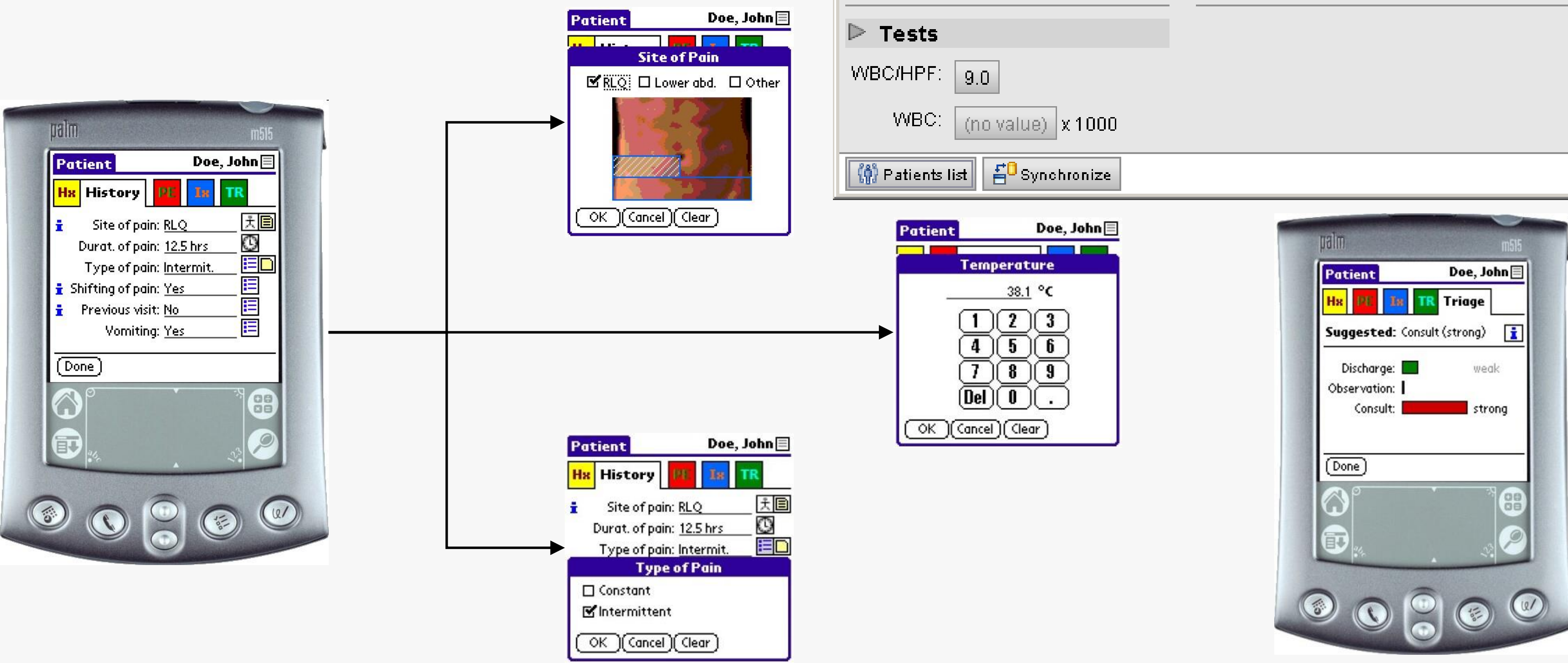
System MET-AP







Met2
Application
MET System – scrotal pain triage
Chang, Carl
ScrotalPain
History
Site of pain: Both Left None 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: 36.6 Celsius
Swelling: Both Left None Right
Tests
WBC/HPF: 9.0
WBC: (no value) x 1000
Patients list
Synchronize



palm
m515
Patient Doe, John
Hx History PE Ix TR
Site of pain: RLQ
Durat. of pain: 12.5 hrs
Type of pain: Intermit.
Shifting of pain: Yes
Previous visit: No
Vomiting: Yes
Done
Site of Pain
RLQ Lower abd. Other
OK Cancel Clear
Type of Pain
Constant
Intermittent
Temperature
38.1 °C
1 2 3 4 5 6 7 8 9 Del 0 .
Triage
Suggested: Consult (strong)
Discharge: weak
Observation:
Consult: strong

Violinmakers competition

Jury's assessment

Sound recording

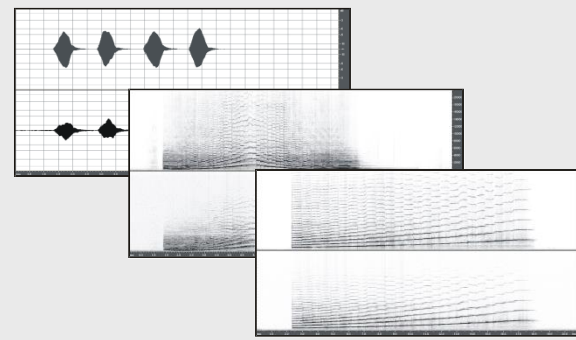
Ranking with respect to:
- **volume of sound (X)**,
- **timbre of sound (Y)**,
- ease of sound emission,
- **equal sound volume of strings (Z)**,
- accuracy of assembly,
- individual qualities

Ranking of violins with respect to **X**

Ranking of violins with respect to **Y**

Ranking of violins with respect to **Z**

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

Ranking according to	Best subset of acoustic features	Number of rules	Ranking fit
volume	A14, E13, D12, G16	62	87%
timbre	E13, D15, G4, G17, D5	99	92%
inter-string equality	D20, D15, A24, D10	64	79%

Summary and conclusions

Summary and conclusions

- Robust Ordinal Regression is a constructive way of learning DM’s preferences.
- It was adapted to three kinds of preference models (value function, outranking relation, decision rules), multiple-criteria ranking, choice and sorting, group decision, hierarchical family of criteria, and decision under risk & uncertainty.

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)

- Robust Ordinal Regression goes along with this recommendation, and as such, it is a representative of the European School of Decision Aiding

Thank you

Acknowledgment to co-operators:

Salvatore Corrente, Salvatore Greco, Miłosz Kadziński,

José Figueira, Jürgen Branke, Vincent Mousseau,…

roman.slowinski@cs.put.poznan.pl