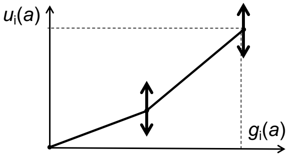
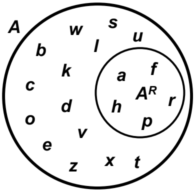
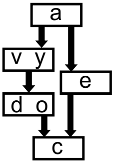






Ordinal regression in the UTA method and robust ordinal regression in the UTA^{GMS} method

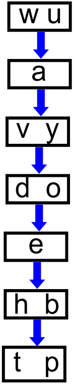
Miłosz Kadziński

Institute of Computing Science
Poznan University of Technology, Poland

Ranking Problem

How to order alternatives from the best to the worst?

ranking



complete

- Imposing order on the set of alternatives according to the Decision Maker's preferences
- Ranking is based on a relative evaluation by comparing alternatives against each other
- **Complete** ranking does not admit incomparability
  - Only preference and indifference are admitted
- Preference implies a better position in the ranking
- Indifference is interpreted as a shared rank
- Complete ranking can be obtained by assigning a score to each alternative
- **Partial** ranking admits incomparability
- Choice and ranking are closely related

- Finite set $A=\{a,b,\ldots\}$ of alternatives
- Consistent family $F=\{g_1,\ldots,g_n\}$ of n criteria; $I=\{1,\ldots,n\}$
- Performance of alternative a on criterion g_i: $g_i(a)$

Illustrative Study

Alternative	$g_1 \uparrow$	$g_2 \uparrow$	$g_3 \uparrow$
Andreev	26	40	44
Brown	2	2	68
Calvet	18	17	14
Dubov	35	62	25
Elmendi	7	55	12
Ferret	25	30	12
Grishuk	9	62	88
Hornet	0	24	73
Ishak	6	15	100
Jope	16	9	0
Kante	26	17	17
Liu	62	43	0

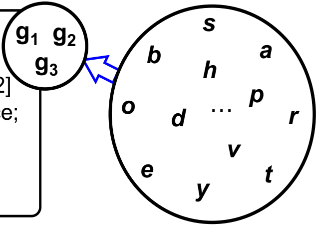
- Medium size firm producing tools for agriculture
- CEO intends to double the production
- **Hire a new sales manager (or a few sales managers)**
- A recruitment agency has interviewed (12) candidates
- Rank the candidates from the best to the worst
- Assign a score (cardinal value) to each candidate

Three criteria:

g_1 : sales management experience; *gain*; sc. [0, 62]

g_2 : international experience; *gain; scale* [2, 62]

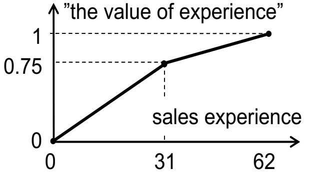
g_3 : human qualities; *gain scale* [0, 100]



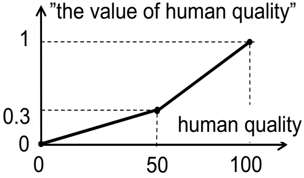
We have the alternatives and performances on all criteria, ordered using the DM's preferences
For simplicity, we assume the scales are delimited by the extreme observed performances

Need for Marginal Value Functions

- **Measurement**, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules
- **Evaluation** is *measuring values* (of Decision Makers)
- We need a (marginal) value function capturing the preference of each alternative on each criterion and the differences of preferences



- The value of 50 is 0.3, and the value of 100 is 1
- Human quality of 100 is far more important than 50
- The difference between 50 and 100 is far more important from the one between 0 and 50
- Compare differences of preferences on one criterion



- *Hypothesis*: the different criteria are separable, and preferences are measurable in terms of differences
- We can compare the differences of preferences on one criterion to the differences of preferences on another one

Additive Value Function

- A **marginal value function** u_i quantifies the preferences on **criterion** g_i
  - *Conjoint interval scale*; the range of preferences, e.g., between 0 and 1
- We need a way to aggregate preferences from different criteria into a comprehensive value: $U(a) = f(u_1(g_1(a)), \ldots, u_n(g_n(a)))$
- *Intuition*: certain criteria are more "important" than others
- *Hypothesis*: preferences on each criterion are independent

An additive value function: a comprehensive value is a weighted sum of marginal values

$$U(a) = \sum_{i=1,\ldots,n} w_i \cdot u_i(g_i(a)) = w_1 \cdot u_1(g_1(a)) + \ldots + w_n \cdot u_n(g_n(a))$$

comprehensive value	weight associated with criterion g_i	marginal value function associated with g_i	performance of alternative *a* on g_i

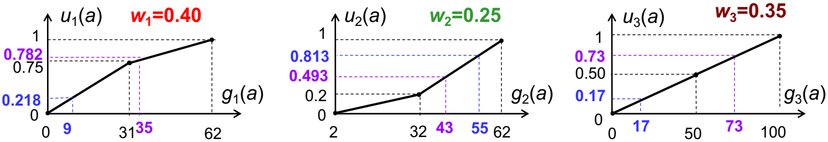
Compensatory model: good values on one criterion can compensate bad values on another criterion

Value function function distinguishes only 2 possible relations between alternatives:

- Preference relation (*P*, >): $a > b \Leftrightarrow U(a) > U(b)$ (asymmetric and transitive)
- Indifference relation (*I*, ~): $a \sim b \Leftrightarrow U(a) = U(b)$ (symmetric, reflexive, and transitive)

Use of Additive Value Function

Assume the marginal value functions and weights are given



	g_1↑	g_2↑	g_3↑
A	26	40	44
B	2	2	68
C	18	17	14
D	35	62	25
E	7	55	12
F	25	30	12
G	9	62	88
H	0	24	73
I	6	15	100
J	16	9	0
K	26	17	17
L	62	43	0

read off marginal values $u_i(a)$

	$u_1(a)$	$u_2(a)$	$u_3(a)$
A	0.629	0.413	0.440
B	0.048	0.000	0.680
C	0.435	0.100	0.140
D	0.782	1.000	0.250
E	0.169	0.813	0.120
F	0.605	0.187	0.120
G	0.218	1.000	0.880
H	0.000	0.147	0.730
I	0.145	0.087	1.000
J	0.387	0.047	0.000
K	0.629	0.100	0.170
L	1.000	0.493	0.000

compute comprehensive values $U(a)$

rank alternatives in decreasing order of $U(a)$

	$U(a)$	*Rank*
A	0.509	4
B	0.257	10
C	0.248	11
D	0.650	1
E	0.313	8
F	0.331	7
G	0.645	2
H	0.292	9
I	0.430	5
J	0.167	12
K	0.336	6
L	0.523	3

COMPREHENSIVE VALUE FOR *A*

$$U(A) = \sum_{i=1,\ldots,n} w_i \cdot u_i(g_i(A)) = w_1 \cdot u_1(26) + w_2 \cdot u_2(40) + w_3 \cdot u_3(44) =$$

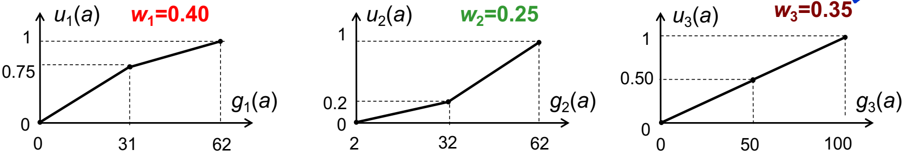
$$= 0.40 \cdot 0.629 + 0.25 \cdot 0.413 + 0.35 \cdot 0.440 = 0.509$$

Additive Value Function

So far, we have assumed that the marginal value functions are normalized in the interval [0,1] and a comprehensive values is defined as a weighted sum of marginal values:

$$U(a) = \sum_{i=1,\ldots,n} w_i \cdot u_i(g_i(a)) = \sum_{i=1,\ldots,n} w_i \cdot u_i(a)$$

- weights w_i are trade-offs between marginal value functions
- they scale contribution of individual criteria in the comprehensive value

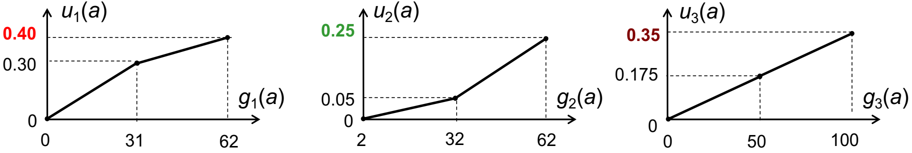


It is possible to incorporate weights into marginal value function by assuming $w_i = u_i(\beta_i)$, for all criteria

- Then, the comprehensive value can be computed as a sum of marginal values

$$U(a) = \sum_{i=1,\ldots,n} u_i(g_i(a)) = \sum_{i=1,\ldots,n} u_i(a)$$

- The criterion's contribution in the comprehensive value is controlled by the maximal marginal value



Need for Easy Preference Information

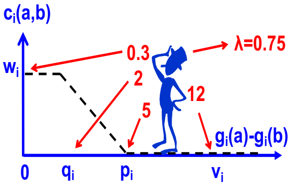
Traditional MCDA methods require **rich (and sometimes) difficult** preference information:

- e.g., many intra-criteria and inter-criteria parameters
- the DM may be overloaded with numbers

They suppose the DM **understands the logic** of a particular model:

- meaning of indifference, preference, and veto thresholds (PROMETHEE, ELECTRE)
- meaning of weights: substitution ratios or relative strengths, ...

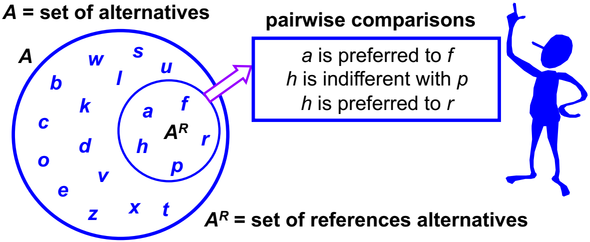
DIRECT PREFERENCE INFORMATION



- The traditional MCDA methods may be **too demanding of cognitive effort** of their users
- We advocate for methods requiring „easy" preference information
- **"Easy" means natural and even partial**

Indirect Preference Information

- Psychologists confirm that DMs are more confident exercising their decisions than explaining them
- The most natural is a **holistic pairwise comparison of some reference alternatives**



A reference set can be composed of:

- **alternatives relatively well known to the DM, especially when *A* is large**
- past decision alternatives on which decision are known
- fictive alternatives, consisting of performances on the criteria, which the DM can easily judge to perform global comparisons

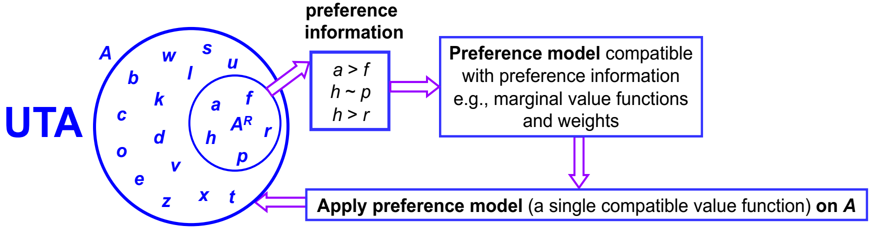
The UTA Method

Disaggregation-aggregation (or ordinal regression) paradigm:

- The holistic preference on a subset $A^R \subseteq A$ is known first, and then a compatible criteria aggregation model (preference model) is inferred from this information to be applied on set A
- Ordinal regression paradigm emphasizes the **discovery of intentions** as an interpretation of actions, rather than as *a priori* position

It is thus concordant with:

- "posterior rationality" principle by March (1978) and "learning from examples" used in AI and knowledge discovery



Traditional **aggregation** paradigm: the preference model is first constructed and then applied on set A to get information about holistic preference

Preference Disaggregation

UTA employs a preference model in the form of an **additive value function** with monotonic marginal value functions u_i

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

The preference model is constructed indirectly based on the DM's holistic judgments:

- In the traditional UTA, a complete ranking of reference alternatives needs to be provided
- It is admitted to provide pairwise comparisons of reference alternatives (partial ranking)

A^R: *a* *f* *h* *r* *p* → holistic preference information → **preference model** compatible with preference information

Constructing preference model via **preference disaggregation**:

- when *a* is judged more preferred (>) than *b*, its comprehensive value should be greater according to the constructed model *U* (value function)
- when *a* and *b* are judged indifferent (~), their comprehensive values should be equal according to the constructed model *U*

$$U(a) > U(b) \Leftrightarrow a > b$$

$$U(a) = U(b) \Leftrightarrow a \sim b$$

Normalization Constraints

Constructing an additive value function via preference disaggregation is subject to the constraints that ensure both the reconstruction of pairwise comparisons and a desired form of the preference model

Normalization constraints:

A marginal value assigned to the worst performance α_i on each criterion g_i needs to be 0:

$$u_i(\alpha_i) = 0,\ i=1,..,n$$

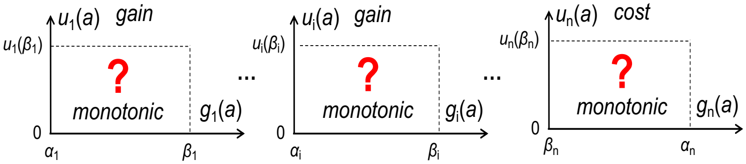
- For gain-type criteria, α_i is the least performance and for cost-type criteria α_i is the greatest performance;

The sum of marginal values corresponding to the performances β_i on all criteria needs to be 1:

$$\sum_{i=1,\dots,n} u_i(\beta_i) = 1$$

- For gain-type criteria, β_i is the greatest performance, and for cost-type criteria, β_i is the least performance;

α_i and β_i can be determined based on the set of all alternatives or as feasible extreme performances delimiting the performance scale, even if they are not observed in the set of alternatives



$$u_i(\alpha_i) = 0 \quad \forall i \in I$$

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

Simple Example (1)

Let us first consider a simplified problem:
- two criteria of gain-type defined over the range [0,10]
- three (reference) alternatives *X*, *Y*, *Z* (see table)
- linear value functions
- reference ranking: *X* > *Y* > *Z*

Alt.	$g_1 \uparrow$	$g_2 \uparrow$
X	10	0
Y	5	5
Z	0	10

• *X*
>
• *Y*
>
• *Z*

Comprehensive value function:

$$U(a)=\sum_{i=1}^{n} u_i\left[g_i(a)\right]=u_1\left[g_1(a)\right]+u_2\left[g_2(a)\right]=\frac{g_1(a)}{\beta_1}u_1(\beta_1)+\frac{g_2(a)}{\beta_2}u_2(\beta_2)$$

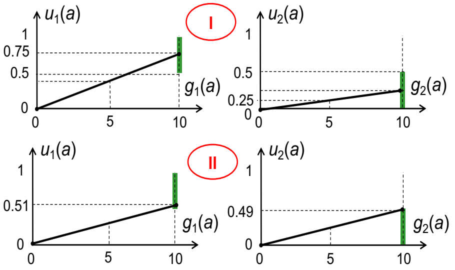
$$U(X)=\frac{10}{10}u_1(10)+\frac{0}{10}u_2(10)=u_1(10)$$

v

$$U(Y)=0.5u_1(10)+0.5u_2(10)$$

v

$$U(Z)=\frac{0}{10}u_1(10)+\frac{10}{10}u_2(10)=u_2(10)$$



Reconstruction works for any:

$$u_1(10)>u_2(10)$$
$$u_1(10)+u_2(10)=1$$

	$u_1(a)$	$u_2(a)$	$U(a)$
X	0.75	0.0	0.75
Y	0.375	0.125	0.5
Z	0.0	0.25	0.25

	$u_1(a)$	$u_2(a)$	$U(a)$
X	0.51	0.0	0.51
Y	0.255	0.245	0.5
Z	0.0	0.49	0.49

Simple Example (2)

Let us first consider a simplified problem:

- two criteria of gain-type defined over the range [0,10]
- three (reference) alternatives *X*, *Y*, *Z* (see table)
- linear value functions
- modified reference ranking: *X* > *Z* > *Y*

Alt.	$g_1 \uparrow$	$g_2 \uparrow$
X	10	0
Y	5	5
Z	0	10

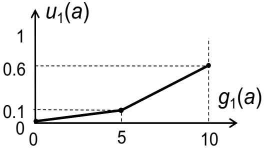
• X > • Y > • Z ⇒ • X > • Z > • Y

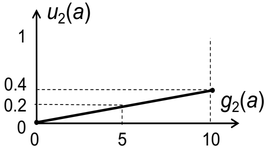
It is not possible to reconstruct the ranking with linear marginal value functions, i.e., to find $u_1(10)$ and $u_2(10)$ satisfying the constrains

$$\boxed{U(X)=u_1(10)} > \boxed{U(Z)=u_2(10)} > \boxed{U(Y)=0.5u_1(10)+0.5u_2(10)}$$

$$\boxed{u_1(10)>u_2(10)>u_1(10)}$$

- One linear piece per each marginal value function u_1, u_2 is not enough
  - Thus, (for some problems) marginal value functions cannot be linear
- Solution: break the marginal value function u_1 in the middle, i.e., consider a function with two linear pieces (additional characteristic point in the middle of the performance scale)





• X > • Z > • Y

Alt.	$u_1(a)$	$u_2(a)$	$U(a)$
X	0.6	0.0	0.6
Y	0.1	0.2	0.3
Z	0.0	0.4	0.4

Piecewise Linear Marginal Value Function

UTA uses **piecewise linear marginal value functions** with the number of linear pieces (characteristic points = breakpoints between these pieces) specified by the DM

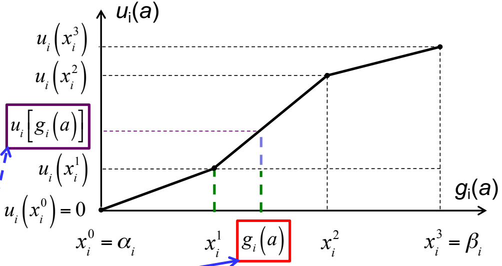
- Linear marginal value function is composed of one linear piece (two characteristic points)
- 2-piecewise linear function is composed of two pieces (three characteristic points)

The intervals [α_i, β_i] are divided, by default) into γ_i **equal sub-intervals** with the end-points ($i \in I$)

$$\boxed{x_i^j = \alpha_i + \frac{j-1}{\gamma_i - 1}\left(\beta_i - \alpha_i\right),\ \ j = 0,\ldots,\gamma_i}$$

- For 2 pieces, the breakpoint is the middle
- For 3 pieces, the breakpoints are in the 1/3 and 2/3 of the performance scale, etc.

In the UTA method, the marginal value of alternative $a \in A$ with performance between the characteristic points is approximated by **linear interpolation**:



$$\boxed{g_i(a) \in \left[x_i^j, x_i^{j+1}\right]}$$

$$\boxed{u_i\left[g_i(a)\right] = u_i\left(x_i^j\right) + \frac{x_i - x_i^j}{x_i^{j+1} - x_i^j}\left[u_i\left(x_i^{j+1}\right) - u_i\left(x_i^j\right)\right]}$$

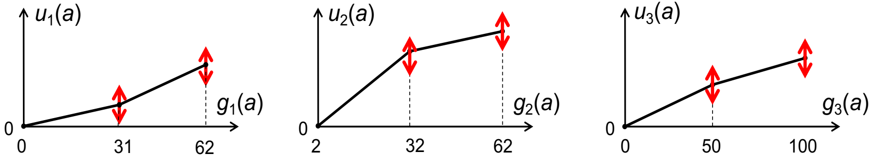
- Start from the marginal value of the worse characteristic point
- Increase it by an appropriate proportion of the differences between marginal values assigned to the two neighboring characteristic points

Monotonicity Constraints

- **The marginal values assigned to the characteristic points are variables in the ordinal regression problem**
- They need to be non-negative for each breakpoint and criterion

$$u_i\left(x_i^j\right) \geq 0, \quad \forall i \text{ and } j$$

- Assume, for our problem, we use **2-piecewise linear marginal values function** for each criterion
  - In general, it is possible to use different numbers of linear pieces for each criterion
- The breakpoints at the ends of perform. scales and precisely in the middle (g_1 – 31, g_2 – 32, g_3 – 62)



- When looking for a value function compatible with the DM's preferences, we need to respect **monotonicity** constraints for marginal value functions
- For each pair of neighboring characteristic points, a marginal value assigned to a better point needs to be at least as high as a marginal value assigned to a worse point

GAIN

$$u_i\left(x_i^{j+1}\right) - u_i\left(x_i^j\right) \geq 0 \quad j = 0,\ldots,\gamma_i\text{-}1$$

COST

$$u_i\left(x_i^j\right) - u_i\left(x_i^{j+1}\right) \geq 0 \quad j = 0,\ldots,\gamma_i\text{-}1$$

Modelling Pairwise Comparisons

- The ordinal regression problem needs to respect the monotonicity and normalization constraints
- The aim is to find the model reproducing the DM's pairwise comparisons using only the model
- When reproducing the pairwise comparisons provided by the DM, we need to be aware that it might not be possible to find a compatible value function
- **Strict inequality can be changed into a non-strict inequality using variable ε**
- To satisfy the constraint, the value of ε should be greater than 0

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

comprehensive value

Preference disaggregation:

- when a > to b, then $U(a) > U(b)$
- when $a \sim b$, then $U(a) = U(b)$

$$\begin{aligned} U(a) &\geq U(b) + \varepsilon \Leftrightarrow a > b \\ U(a) &= U(b) \qquad \Leftrightarrow a \sim b \end{aligned}$$

Note: we present a simplified variant of UTA that will make the presentation of other methods (UTA^{GMS}) more straightforward

Ordinal Regression

The marginal value functions (breakpoint variables) are estimated by **solving the LP problem**

Maximize the epsilon in the strict preference comparisons of reference alternatives:

- to reproduce the DM's pairwise comparisons (preference and indifference judgments),
- while respecting the normalization, monotonicity, and non-negativity constraints.

Maximize ε **objective function = maximize epsilon**

subject to

$$\left.\left.\begin{array}{ll} U(a) \geq U(b) + \varepsilon \Leftrightarrow a > b & \forall a,b \in A^R \\ U(c) = U(d) \quad \Leftrightarrow c \sim d & \forall c,d \in A^R \end{array}\right\} R \quad \right\} C(R)$$

reference ranking

$$\left.\begin{array}{ll} \sum_{i=1}^{n} u_i(\beta_i) = 1 & \\ u_i(\alpha_i) = 0, & \forall i \in I \\ u_i\left(x_i^{j+1}\right) - u_i\left(x_i^j\right) \geq 0, & j = 0, \ldots, \gamma_i - 1;\ \forall i \in I \\ u_i\left(x_i^j\right) \geq 0,\ \forall\, i \text{ and } j & \end{array}\right\} C$$

normalization

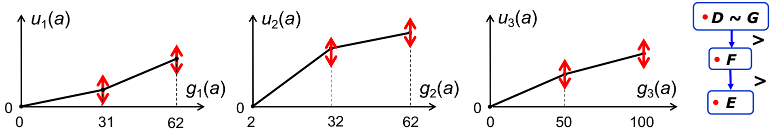
monotonicity

non-negativity

Y. Siskos, E. Grigoroudis, N. Matsatsinis, UTA Methods. In: S. Greco, M. Ehrgott, J. Figueira (Eds.), *State of the Art in Multiple Criteria Decision Analysis*, Springer, pp. 315–365, 2016.

Ordinal Regression - Example

- For our problem, let us select a subset of reference alternatives $A^R = \{D, E, F, G\}$
- A complete DM's reference ranking: **Dubov ~ Grishuk > Ferret > Elmendi**
- Marginal value functions with two pieces for each criterion



Maximize ε

subject to

$$u_1(35) + u_2(62) + u_3(25) = u_1(9) + u_2(62) + u_3(88)$$
$$u_1(9) + u_2(62) + u_3(88) \geq u_1(25) + u_2(30) + u_3(12) + \varepsilon$$
$$u_1(25) + u_2(30) + u_3(12) \geq u_1(7) + u_2(55) + u_3(12) + \varepsilon$$

reference ranking

$$u_1(62) + u_2(62) + u_3(100) = 1$$
$$u_1(0) = 0, \quad u_2(2) = 0, \quad u_3(0) = 0$$

normalization

C(*R*)

$$u_1(31) \geq u_1(0), u_1(62) \geq u_1(31)$$
$$u_2(32) \geq u_2(2), u_2(62) \geq u_2(32)$$
$$u_3(100) \geq u_3(50), u_3(50) \geq u_3(0)$$

monotonicity

$$u_1(0), u_1(31), u_1(62), u_2(2), u_2(32), u_2(62),$$
$$u_3(0), u_3(50), u_3(100) \geq 0$$

non-negativity

Alt.	$g_1 \uparrow$	$g_2 \uparrow$	$g_3 \uparrow$
Dubov	35	62	25
Elmendi	7	55	12
Ferret	25	30	12
Grishuk	9	62	88

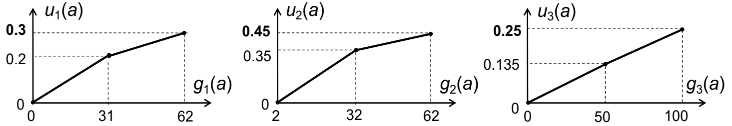
Let us denote the optimal solution of the ordinal regression problem by ε^*

Compatible Value Function (1)

- **If $\varepsilon^*>0$ and $C(R)$ is feasible**, then the polyhedron of feasible solutions for $u_i(a)$ is not empty, and there exists at least one v. f. $U(a)$ compatible with DM's reference ranking on A^R
- Let us inspect an example, compatible model

Alt.	$g_1\uparrow$	$g_2\uparrow$	$g_3\uparrow$
D	35	62	25
E	7	55	12
F	25	30	12
G	9	62	88

• ***D ~ G***
>
• ***F***
>
• ***E***



	$u_1(a)$	$u_2(a)$	$u_3(a)$	$U(a)$	*Rank*
A	0.168	0.377	0.119	0.663	4
B	0.013	0.000	0.176	0.189	11
C	0.116	0.175	0.038	0.329	10
D	0.213	0.450	0.068	0.730	1
E	0.045	0.427	0.032	0.504	6
F	0.161	0.327	0.032	0.520	5
G	0.058	0.450	0.222	0.730	1
H	0.000	0.257	0.188	0.445	7
I	0.039	0.152	0.250	0.440	8
J	0.103	0.082	0.000	0.185	12
K	0.168	0.175	0.046	0.389	9
L	0.300	0.387	0.000	0.687	3

Comprehensive values of references alternatives:

$U(D) = u_1(35) + u_2(62) + u_3(25) = 0.730$

$U(G) = u_1(9) + u_2(62) + u_3(88) = 0.730$

$U(F) = u_1(25) + u_2(30) + u_3(12) = 0.520$

$U(E) = u_1(7) + u_2(55) + u_3(12) = 0.504$

The reference ranking D ~ $G > F > E$ is reproduced because:

$$U(D) = U(G) > U(F) > U(E)$$

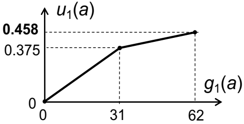
Somehow, by the way, i.e., by applying the inferred value function, we can order all alternatives (see ***Rank*** column), including the non-reference ones, from the best to the worst

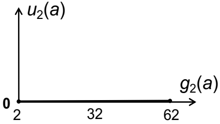
Compatible Value Function (2)

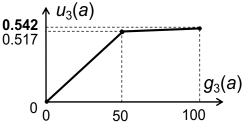
- Let us inspect another solution (e.g., corresponding to max ε^*)
- Increased maximal share in the comprehensive value of criterion g_1 and g_3 (the shapes are more concave) and zeroed maximal share of criteria g_2

Alt.	$g_1\uparrow$	$g_2\uparrow$	$g_3\uparrow$
D	35	62	25
E	7	55	12
F	25	30	12
G	9	62	88

• ***D ~ G***
>
• ***F***
>
• ***E***







	$u_1(a)$	$u_2(a)$	$u_3(a)$	$U(a)$	Rank
A	0.315	0.000	0.456	0.771	1
B	0.024	0.000	0.527	0.551	5
C	0.218	0.000	0.145	0.363	10
D	0.386	0.000	0.259	0.645	2
E	0.085	0.000	0.124	0.209	11
F	0.303	0.000	0.124	0.427	9
G	0.109	0.000	0.536	0.645	2
H	0.000	0.000	0.529	0.529	6
I	0.073	0.000	0.542	0.614	4
J	0.194	0.000	0.000	0.194	12
K	0.315	0.000	0.176	0.491	7
L	0.458	0.000	0.000	0.458	8

Comprehensive values of references alternatives:

$U(D) = u_1(35) + u_2(62) + u_3(25) = 0.645$

$U(G) = u_1(9) + u_2(62) + u_3(88) = 0.645$

$U(F) = u_1(25) + u_2(30) + u_3(12) = 0.427$

$U(E) = u_1(7) + u_2(55) + u_3(12) = 0.209$

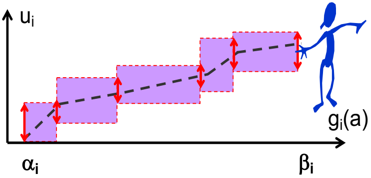
The reference ranking D ~ $G > F > E$ is reproduced because:

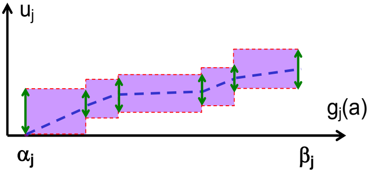
$$U(D) = U(G) > U(F) > U(E)$$

There might be some changes in the ranking depending on the selected compatible model

Selection of Single Compatible Value Function

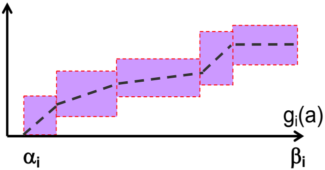
Interactive selection of a single value function within bounds ensuring the reproduction of DM's pairwise comparisons, supported by the graphical interface of decision support system





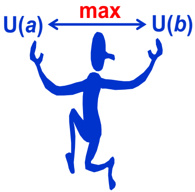
Applying some pre-defined rule (algorithm) for selecting a single (representative) value function

"central value function"



"the most discriminant value function"

for all pairs (a,b) such that a is preferred to b by the DM, find such a value function that maximizes the difference between $U(a)$ and $U(b)$



The UTA Method – Summary

Preference information

- Indirect: a complete preorder on a subset of reference alternatives A^R (can be generalized to pairwise comparisons)
- The number of character points γ_i (linear pieces) for each marginal value function

UTA

Preference model:

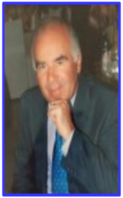
- Additive value function with piecewise linear marginal value functions

Technique

- Linear programming technique solving ordinal regression problem, to infer a value function so that the ranking obtained through its application on A^R is (as) consistent (as possible) with the reference one

Recommendation:

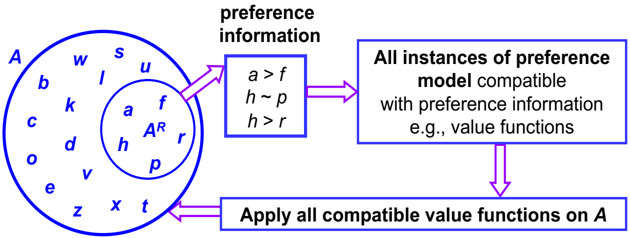
- A complete ranking of all alternatives (ranks and scores = comprehensive values)



E. Jacquet-Lagreze, Y. Siskos, Assessing a set of aadditive utility function for multicriteria decision-making, the UTA methods, *European Journal of Opeartional Research, 101:151-164,* 1982.

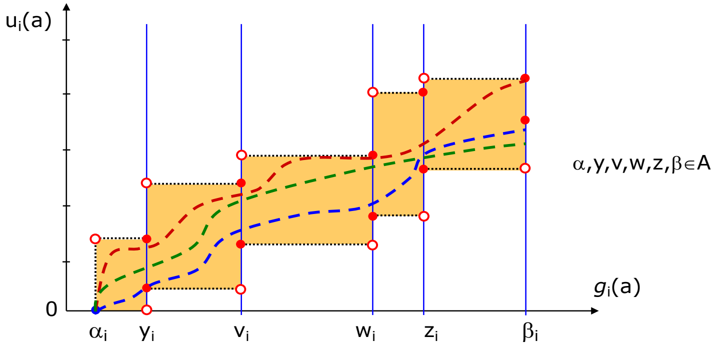
- **Question**: what is the consequence of using all compatible preference models on set *A*?
- **Robustness analysis** verifies whether the conclusions are valid for all or most plausible sets of value in the model
- Instead of arbitrarily selecting a single compatible value model, we will use all of them on set A
- If the set of compatible value functions is non-empty, there are, typically, infinitely many such functions
- **Robust ordinal regression** is the robust variant of the ordinal regression paradigm

UTA^{GMS}



A Set of Compatible Value Functions

- Linear or piece-wise linear marginal value functions are arbitrary
- **General marginal value function** $u_i(a)$ with characteristic points fixed on all actual performances of alternatives from set A



- **Marginal values** in characteristic points are **unknown**
- In fact, they are intervals - all compatible functions are considered
- In the area, the compatible marginal functions must be **monotonic**

Verifying Compatibility

- The preference information is a **partial preorder** on a subset of reference alternatives $A^R \subseteq A$
- $B^R \subseteq A^R \times A^R$ is the set of **pairs of reference alternatives compared by the DM**
- A value function is called **compatible** if it is able to restore all pairwise comparisons from B^R

Maximize ε — objective function = maximize epsilon

subject to

$$\left.\begin{array}{l}\left.\begin{array}{ll} U(a) \geq U(b) + \varepsilon \Leftrightarrow a > b \\ U(a) = U(b) \quad \Leftrightarrow a \sim b \end{array}\right\} \forall (a,b) \in B^R \end{array}\right\} R$$

pairwise comparisons

$$\left.\begin{array}{l} \sum_{i=1}^{n} u_i(\beta_i)=1 \\ u_i(\alpha_i) = 0, \quad \forall i \in I \\ u_i\left(x_i^{j+1}\right) - u_i\left(x_i^j\right) \geq 0, j = 1, \ldots, n_i(A) - 1;\ \forall i \in I \\ u_i(\beta_i) \geq u_i\left(x_i^j\right) \geq u_i(\alpha_i),\ \forall\ i \text{ and } j \end{array}\right\} C$$

R and C together: $C(R)$

normalization

monotonicity

bounds (non-negativity)

where $x_i^1, x_i^2, \ldots, x_i^{n_i(A)}$ are ordered performances of all alternatives in A such that $x_i^j < x_i^{j+1}$ for $j = 1, \ldots, n_i(A) - 1$ and $n_i(A) \leq n$ and $\alpha_i \leq x_i^1$ and $x_i^{n_i(A)} \leq \beta_i$

The Necessary and the Possible

When exploiting a set of all compatible value functions U^R, we are interested in **two preference relations** verified for all pairs of alternatives $(a,b) \in A \times A$:

- ***a* is necessarily preferred to *b*** ($a \succsim^N b$) iff $U(a) \geq U(b)$ **for all** compatible value functions
- $\succsim^N$ means necessary preference relation
  - the minimal difference $d(a,b) = U(a) - U(b)$ in set U^R is greater than or equal to 0
  - $U(b) > U(a)$ is not possible for any function in U^R

N

- ***a* is possibly preferred to *b* ($a \succsim^P b$)** iff $U(a) \geq U(b)$ **for at least one** compatible value function
- $\succsim^P$ means possible preference relation
  - the maximal difference $D(a,b) = U(a) - U(b)$ in set U^R is greater than or equal to 0
  - $U(a) \geq U(b)$ is possible for some function in U^R

P

It is impossible to infer one relation from another because necessary and possible relations **are not dual** (it is possible that $a \succsim^N b$ and still $b \succsim^P a$)

Verifying the Truth of the Necessary and Possible Relations

A set of all compatible value functions U^R is defined by constraint set $C(R)$

Given a pair of alternatives $a,b \in A$, there **exist two sensible ways for verifying the truth of $\succsim^P$ and $\succsim^N$**

The possible relation:

- $a \succsim^P b \Leftrightarrow D(a,b) \geq 0$ where $\boldsymbol{D(a,b) = \text{Max}\,[U(a)-U(b)]}$ subject to $C(R)$ (ε set to a small positive value)

  $\boldsymbol{D(a,b) \geq 0}$ means that **for at least one** compatible value function a is at least as good as b

- **Assume the truth of $a \succsim b$ and prove it holds in set U^R**

  $a \succsim^P b \Leftrightarrow$ if $\varepsilon^* \geq 0$ where $\varepsilon^* = \text{Max}\ \varepsilon$, subject to $C(R) \cup [U(a) \geq U(b)]$ and constraint set is feasible

The necessary relation:

- $a \succsim^N b \Leftrightarrow d(a,b) \geq 0$ where $\boldsymbol{d(a,b) = \text{Min}\,[U(a)-U(b)]}$ subject to $C(R)$ (ε set to a small positive value)

  $\boldsymbol{d(a,b) \geq 0}$ means that **for all** compatible value functions a is at least as good as b

- **Assume the truth of $b \succ a$** (i.e., the inverse relation) **and prove it does not hold in set U^R**

  $a \succsim^N b \Leftrightarrow$ if $\varepsilon^* \leq 0$ where $\varepsilon^* = \text{Max}\ \varepsilon$, subject to $C(R) \cup [U(b) \geq U(a) + \varepsilon]$ or constraint set is infeasible

Some properties:

N AND P

- $a \succsim^N b \Rightarrow a \succsim^P b$ (the necessary implies the possible)
- $\succsim^N$ is a **partial preorder** (i.e., $\succsim^N$ is reflexive and **transitive**)
- $\succsim^P$ is **strongly complete** (i.e., for all $a,b \in A$, $a \succsim^P b$ or $b \succsim^P a$)
  and **negatively transitive** (i.e., for all $a,b,c \in A$, *not* $a \succsim^P b$ and *not* $b \succsim^P c \Rightarrow$ *not* $a \succsim^P c$),
  (in general, $\succsim^P$ is not transitive)

Impact of **pairwise comparisons**: $a \succsim b \Rightarrow a \succsim^N b$ and $a \succ b \Rightarrow$ *not* $b \succsim^P a$

In the **absence of any preference information**:

- necessary relation boils down to weak dominance relation
- possible relation is a complete relation

For **complete pairwise comparisons** (complete preorder in A):

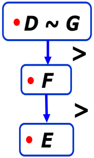
- necessary relation = possible relation

The Necessary Preference – First Iteration

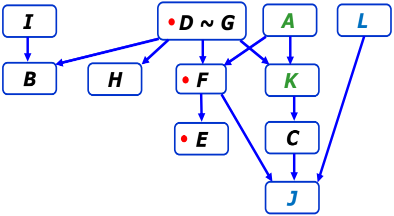
- **Necessary relation** can be presented using a Hasse diagram
- **Hasse diagram** is a type of mathematical diagram used to represent a finite partially ordered set in the form of a drawing of its **transitive reduction** ($A \succsim^N K$ and $K \succsim^N C \Rightarrow A \succsim^N C$)
- The necessary relation includes the preference information, the dominance relation, and consequences of applying U^R

Preference information

Alt.	$g_1 \uparrow$	$g_2 \uparrow$	$g_3 \uparrow$
D	35	62	25
E	7	55	12
F	25	30	12
G	9	62	88



NECESSARY RANKING



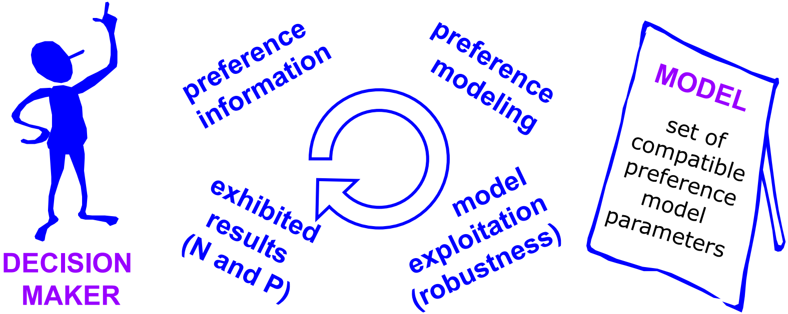
Example dominance relations

Alt.	$g_1 \uparrow$	$g_2 \uparrow$	$g_3 \uparrow$
A	26	40	44
K	26	17	17
L	62	43	0
J	16	9	0

- Pairs such as (*I*,*B*), (*D*,*H*), and (*A*,*E*) are compared in the same way by all compatible value functions
- For pairs such as (*I*,*D*), (*H*,*F*), and (*L*,*K*), the results of a comparison is not univocal

mutual learning of the model and the Decision Maker

the model learns preferences of the Decision Maker



the DM learns from the consequences of applying the model

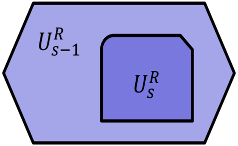
Incremental Set of Pairwise Comparisons

Reference pairwise comparisons in growing sets: $B_1^R \subset B_2^R \subset \ldots \subset B_{t-1}^R \subset B_t^R$

- The reference ranking of alternatives from B_s^R does not change in B_{s-1}^R, $s = 2, \ldots, t$
- It makes sense only to compare alternatives not related by the necessary preference

Each set B_s^R, $s = 1, \ldots, t$, corresponds to a set of compatible value functions U_s^R, the necessary relation $\succsim_s^N$ and the possible relation $\succsim_s^P$

- Each time we pass from B_{s-1}^R to B_s^R, we add to $C(R)$ new constraints
- Nested set of compatible value functions: $U_1^R \supseteq U_2^R \supseteq \ldots \supseteq U_{t-1}^R \supseteq U_t^R$
- **The set of compatible value functions becomes narrower** (and certainly not wider)

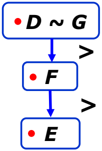


- **The necessary relation is enriched**: $\succsim_{s-1}^N \subseteq \succsim_s^N$
- What is certainly true becomes richer
- **The possible relation is impoverished**: $\succsim_{s-1}^P \supseteq \succsim_s^P$
- What is possibly true becomes narrower

The Necessary Preference – Secod Iteration

1st iteration

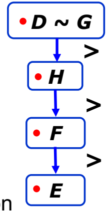
initial preference information



Reference alternatives

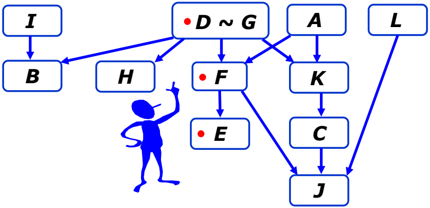
Alt.	$g_1 \uparrow$	$g_2 \uparrow$	$g_3 \uparrow$
D	35	62	25
E	7	55	12
F	25	30	12
G	9	62	88
H	0	24	73

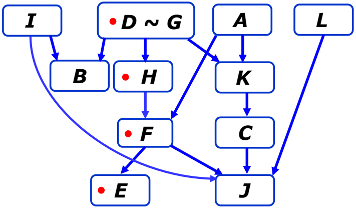
2nd iteration



enriched preference information (*H* preferred to *F*)

H and *F* are incomparable in terms of the necessary relation





- The necessary relation is enriched; see, e.g., (*H*,*F*), (*H*,*E*), and (*I*,*J*)
- The interaction should be continued until the necessary ranking is decisive enough
- There exist active learning algorithms for suggesting pairs to be compared next

The UTA^{GMS} Method – Summary

UTA^{GMS}

Preference information

- Indirect: a partial preorder (i.e., pairwise comparisons) on a subset of reference alternatives A^R

Preference model:

- A set of additive value functions with general marginal value functions (can be – and more often is – used with linear or piecewise linear functions)

Technique

- Linear programming technique solving ordinal regression problem to infer a set of all compatible value functions and to determine the necessary and possible preference relations

Recommendation:

- An incomplete (necessary) ranking of all alternatives; the possible relation

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

The GRIP Method

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 alternatives, e.g., „*x* is preferred to *y* at least as much as *w* is preferred to *z*" which is represented by the LP constraint:

$$U(x)-U(y) \geq U(w)-U(z)$$

- ***partial comparisons of intensities of preference*** between some pairs of reference alternatives on particular criteria, e.g., „*x* is preferred to *y* more than *w* is preferred to *z*, on criterion g_i", which is represented by the LP constraint:

$$u_i(x)-u_i(y) > u_i(w)-u_i(x)$$

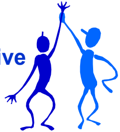
GRIP can handle other kinds of preference information, like **local tradeoffs**, e.g.:

- „the increase of $g_i(a)$=6 by 3 is at least as attractive as the decrease of $g_j(a)$=8 by 2", which is represented by the LP constraint: $u_i(9)-u_i(6) \geq u_j(8)-u_j(6)$

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. Oper. Res.*, 195, 460-486, *2009.*

The Most Representative Value Function

- The principle of the most representative value function U^{REP} approach is:
  - **one for all**: one value function represents all compatible value functions
  - **all for one**: all compatible value functions contribute to the definition of the most representative value function



- The idea is to select among all compatible value functions the "most discriminant" value function for consecutive alternatives in the necessary ranking, i.e.:

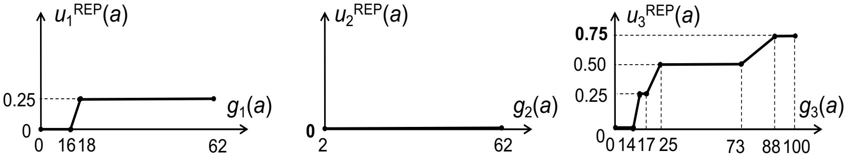
  that value function which **maximizes** the difference of scores between alternatives (a,b) **related** by the necessary preference ($a \succsim^N b$ but (*not* $b \succsim^N a$))

  add the following constraints to the LP constraints: $U(a) \geq U(b)+\varepsilon$

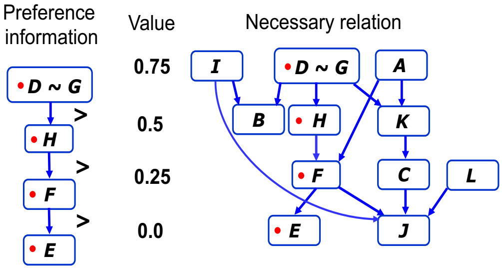
- To tie-breaking, one may **minimize** the difference of scores between alternatives (c,d) **not related** by the necessary preference ((*not* $c \succsim^N d$) and (*not* $d \succsim^N c$))

  add the following constraints to the LP constraints: $U(c)-U(d) \leq \delta$ and $U(d)-U(c) \leq \delta$

- **Maximize** $M\varepsilon - \delta$, where M is a "big value"



	$u_1(a)$	$u_2(a)$	$u_3(a)$	$U^{REP}(a)$	*Rank*
A	0.25	0	0.5	0.75	1
B	0	0	0.5	0.5	5
C	0.25	0	0	0.25	8
D	0.25	0	0.5	0.75	1
E	0	0	0	0	11
F	0.25	0	0	0.25	8
G	0	0	0.75	0.75	1
H	0	0	0.5	0.5	5
I	0	0	0.75	0.75	1
J	0	0	0	0	11
K	0.25	0	0.25	0.5	5
L	0.25	0	0	0.25	8



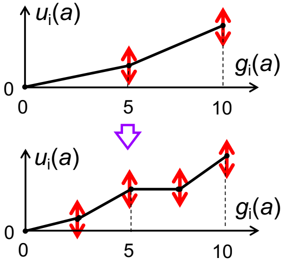
- Allows understanding of the necessary ranking
- Flattens the consequences of applying all compatible value functions

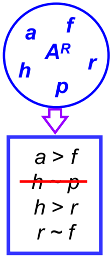
Dealing with Incompatibility

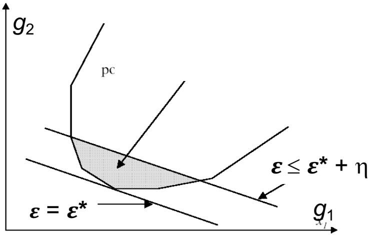
If there is no value function $U(a)$ compatible with the DM's preference information

Three possible moves:

- **increasing the preference model's flexibility** (e.g., increase the number of linear pieces γ_i for u_i or add interactions between criteria)
- post optimal **search for the best function(s)** having a sub-minimal error that is consistent with the DM's preference information to a large extend
- **revision of the preference information** on A^R – accept that some mistake has been made; eliminate or change some pairwise comparisons (or part of the complete ranking)







Resolving Inconsistency (1)

- Restoring consistency relies on finding a **minimal subset of constraints that need to be removed** from the constraint set
- In the context of UTA and UTA^{GMS}, some constraints cannot be eliminated (constraint set *C* composed of the normalization, monotonicity, and non-negativity constraints)
- Constraints related to each pairwise comparison become a candidate for being eliminated
- We associate a **binary variable** (taking a value of either 0 or 1) with each pairwise comparison
- Rewrite each constraint, in a specific way, using a binary variable

How to use the binary variables?

preference

$$U(a) > U(b)$$

$$\Downarrow$$

$$U(a) > U(b) - v_{a,b}$$

come back to the original constraint

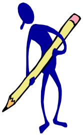
$$v_{a,b} = 0 \Rightarrow U(a) > U(b)$$

$$v_{a,b} = 1 \Rightarrow U(a) > U(b) - 1$$

constraint always satisfied

indifference

$$U(c) = U(d) \Rightarrow \begin{array}{l} U(c) \geq U(d) \\ U(d) \geq U(c) \end{array}$$



$$\Downarrow$$

$$U(c) \geq U(d) - v_{c,d}$$

$$U(d) \geq U(c) - v_{c,d}$$

Resolving Inconsistency (2)

- Minimize the cardinality of the subset of pairwise comparisons to be removed to restore consistency
- Minimize the sum of binary variables, each associated with constraint(s) translating a specific comparison

$$\text{Min} \rightarrow \text{V} = \sum_{(a,b)\in A^R x A^R} v_{a,b} + \sum_{(c,d)\in A^R x A^R} v_{c,d}$$

subject to

$$\left.\begin{array}{l} U(a) > U(b) - v_{a,b} \Leftrightarrow a > b \\ U(c) \geq U(d) - v_{c,d} \Leftrightarrow c \sim \text{d} \\ U(d) \geq U(c) - v_{c,d} \Leftrightarrow c \sim \text{d} \end{array}\right\} \begin{array}{l} \forall a,b \in A^R \\ \\ \forall c,d \in A^R \end{array}$$

$$v_{a,b},\ v_{c,d} \in \{0,1\},\ \forall (a,b),\ (c,d) \in A^R \text{x} A^R$$

$$C \text{ (normalization, monotonicity, non-negativity)}$$

- Assume the optimal value of the objective function is V^*
- The binary variables v, which are equal to 1, correspond to the constraints (pairwise comparisons) that need to be removed to restore consistency
- For example, $V^* = 2$ with $v_{D,C} = 1$ and $v_{B,E} = 1$ imply that one needs to remove two statements: $D > C$ and $B \sim E$

$F > E$, $D > C$, $A > J$, $B \sim E$, $F \sim G$

$$\text{Min} \rightarrow \text{V} = v_{F,E} + v_{D,C} + v_{A,J} + v_{B,E} + v_{F,G}$$

$$\left.\begin{array}{l} U(F) > U(E) - v_{F,E} \\ U(D) > U(C) - v_{D,C} \\ U(A) > U(J) - v_{A,J} \\ U(B) \geq U(E) - v_{B,E} \\ U(E) \geq U(B) - v_{B,E} \\ U(F) \geq U(G) - v_{F,G} \\ U(G) \geq U(F) - v_{F,G} \\ v_{F,E},\ v_{D,C},\ v_{A,J},\ v_{B,E},\ v_{F,G} \in \{0,1\} \\ C \text{ (norm, monot, non-negat)} \end{array}\right\}$$

constraints related to 5 pairwise comparisons rewritten using the binary variables

Resolving Inconsistency (3)

- There may exist more minimal subsets of constraints and corresponding pairwise comparisons underlying inconsistency
- Present all subsets underlying inconsistency to the DMs so that they decide which should be eliminated

resolving inconsistency in preference information *PI*

$PI_1 \subset PI$ $PI_2 \subset PI$ $\cdots$ $PI_K \subset PI$

- The optimal (minimal) value of the objective function in the k-th iteration: V_k^*
- The subset of constraints to identified in the k-th iteration and corresponding to Pi_k

$$\boxed{V_k = \{v_{a,b} \text{ such that } (a,b) \in A^R x A^R \text{ and } v_{a,b} = 1\}}$$

- To identify other subsets underlying inconsistency, solve the problem considered in the previous iteration while forbidding finding again the same solution
- The sum of V_k^* binary variables set to 1 in the previous iteration needs to be less or equal to V_k^* - 1

$$\boxed{\sum_{v_{a,b} \in V_k} v_{a,b} \leq V_k^* - 1}$$

- In the first iteration, $V_1^* = 2$ with $v_{D,C} = 1$ and $v_{B,E} = 1$ imply that one needs to remove statements: $D > C$ and $B \sim E$

$$\boxed{v_{D,C} + v_{B,E} \leq 2 - 1}$$

- If there exist other subsets underlying inconsistency, they will be identified; if not, infeasibility
- $V_2^* = 2$ with $v_{D,C} = 1$ and $v_{F,G} = 1$ imply one needs to remove statements: $D > C$ and $F \sim G$

UTA-like Methods - Summary

- Techniques for building value functions are **prevailing in MCDA**
  - Intuitive interpretation and easy computation
- Preference disaggregation: inferring an analytical preference model consistent with the DM's indirect, holistic preferences ("posterior rationality", "learning from examples")
  - **Preference disaggregation** has been one of the most important methodological streams in MCDA in the 21^{st} century
  - Close links to data mining and statistical machine learning (identifying patterns, extracting knowledge from data)

- The acceptance of such a preference model is accomplished through a repetitive interaction between the model and the DM.
- The UTA-like methods are **widely used** in, e.g., financial management (portfolio selection, business financing, country risk assessment), energy management and planning, marketing (new products, consumer behavior, sales strategy), environmental management, project evaluation, facility location

ROR Methods - Example Applications

- **Ranking wastewater infrastructure alternatives** in Switzerland (Zheng and Lienert, 2018)
- **Siting an urban waste landfill** in Italy (Angilella et al., 2016)
- Global **e-government evaluation** (the use of ICT technologies to provide digital services to citizens and businesses) (Siskos et al., 2014)
- Ranking **therapeutic categories** (antibiotics, gastrointestinal, etc.) for a multinational company to effectively form its investment strategy in the pharmaceutical market of Greece (Mastorakis and Siskos, 2015)
- Sorting **silver nanoparticles synthesis protocols** to risk categories by US Environmental Protection Agency (Kadziński et al., 2018)
- Classification of **activities to be outsourced** in the civil constriction of a brewery in Brazil (Palha et al., 2016)