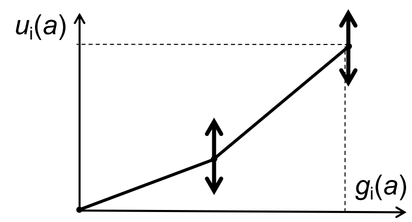
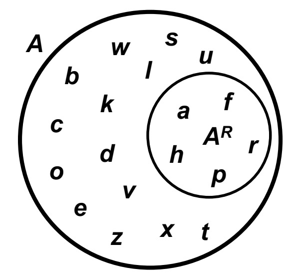
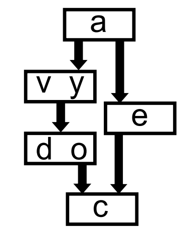






Ordinal regression in the UTA method and robust ordinal regression in the UTA^{GMS} method

Miłosz Kadziński

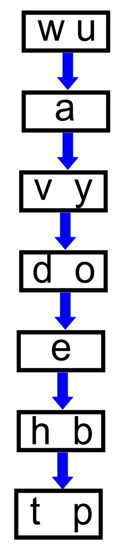
Institute of Computing Science
Poznan University of Technology, Poland

[1] Ten wykład będzie poświęcony problemom porządkowania z wykorzystaniem miar jakości wariantów. Sposób ich obliczenia będzie zupełnie inny niż w metodach opartych na relacji przewyższania. Skupimy się na wykorzystaniu modelu opartego na funkcji wartości, która tłumaczy oceny na każdym kryterium do wartości cząstkowych, a następnie agreguje je do wartości całkowitych (globalnych). Położymy nacisk na metody wykorzystujące pośrednią informację preferencyjną w postaci porównań parami. Rozpoczniemy od omówienia, a dla niektórych z was przypomnienia, metody UTA, która wykorzystuje ranking wariantów referencyjnych do konstrukcji pojedynczej funkcji wartości. UTA implementuje tzw. zasadę regresji porządkowej. Głównym bohaterem wykładu będzie jednak odporna regresja porządkowa na przykładzie metody UTA-GMS. Postuluje ona wykorzystanie wszystkich funkcji wartości spójnych z preferencjami decydenta oraz przeprowadzenie analizy odporności, a więc badania wpływu wykorzystania takich funkcji na stabilność rekomendowanego wyniku.

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: $g_i(a)$

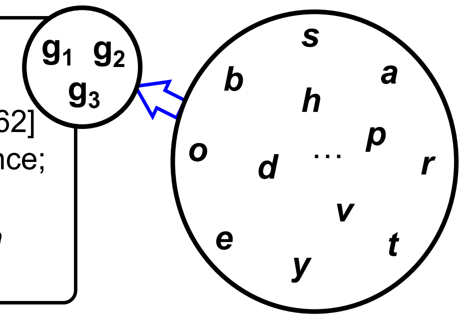
[2] Zacznijmy od przypomnienia istoty problemów porządkowania. Celem jest tu uszeregowanie wariantów od najlepszego do najgorszego. Dziś pokażemy, jak konstruować ranking zupełny, który dopuszcza jedynie relacje preferencji i nierozróżnialności. Oczywiście, preferencja dla jednego z wariantów oznacza wyższą pozycję w rankingu, natomiast nierozróżnialność jest równoznaczna z dzieleniem tej samej pozycji. Omówimy metodę stosującą ocenę punktową dla każdego z wariantów. W ten sposób ranking może być postrzegany jako liczbowy lub kardynalny. W drugiej części wykładu skupimy się na konstrukcji rankingu częściowego, który dopuszcza nieporównywalność. Ranking ten będzie miał charakter porządkowy, bo o umiejscowieniu w nim wariantów nie będą już decydować pojedyncze liczby.

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]

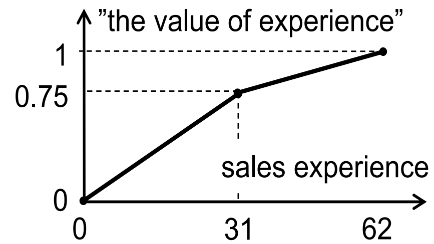


We have the alternatives and performances on all criteria, ordered using the DM's preferences
For simplicity, we assume the scales are delimited by the extreme observed performances

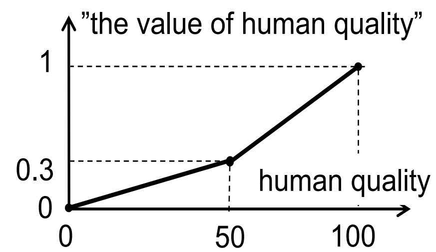
[3] Nasz problem ilustrujący dotyczy średniej wielkości firmy, która produkuje narzędzia dla rolnictwa. Jej dyrektor zamierza podwoić produkcję w ciągu najbliższych dwóch lat. W tym celu musi zatrudnić nowego kierownika sprzedaży. Agencja rekrutacyjna przeprowadziła rozmowy z kandydatami na to stanowisko i oceniła ich pod względem trzech istotnych kryteriów typu zysk. Rozważymy doświadczenie w zarządzaniu sprzedażą, doświadczenie międzynarodowe oraz cechy osobowościowe. Oceny na każdym kryterium mają charakter punktowy, a dla ułatwienia przyjmujemy, że skrajne obserwowalne wartości ograniczają skalę ocen. Przykładowo, na pierwszym kryterium najmniejsza ocena to 0, a największa 62. Na tym etapie dysponujemy wariantami i ich ocenami na poszczególnych kryteriach. Celem jest uszeregowanie dwunastu kandydatów od najlepszego do najgorszego. Ich nazwiska celowo zaczynają się od liter od A do L tak by potem można się do nich odwoływać skrótowo.

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

[4] Metody, które omówimy, opierają się na pojęciach pomiaru i oceny. Pomiar definiuje się jako przypisanie liczb obiektom lub zdarzeniom zgodnie z pewnymi z góry określonymi regułami, natomiast ocena to pomiar wartości. My skupimy się na pomiarze wartości, które wariantom przypisują decydenci. W tym celu wykorzystamy narzędzie zwane cząstkową funkcją wartości. Jej rolą dla danego kryterium jest uchwycenie preferencji każdego z wariantów oraz różnic preferencji. Na slajdzie widzicie przykładową cząstkową funkcję wartości dla kryterium cech osobowościowych. Korzystając z niej, można się dowiedzieć, że jakość dla oceny 50 to 0.3 i jest ona znacznie mniejsza od jakości dla oceny 100-punktowej. Ponadto, można zauważyć, że różnica preferencji między 0 i 50 punktów jest zdecydowanie mniejsza niż różnica między 50 a 100 punktów. Oznacza to, że decydent wysoko ceni korzystne cechy osobowościowe kandydatów. Takie cząstkowe funkcje możemy rozpatrywać też dla innych kryteriów. Ta w środkowej części slajdu pokazuje, jak określić wartość doświadczenia sprzedaży. Zakładając, że kryteria są niezależne, możemy porównywać różnice preferencji między kryteriami.

Additive Value Function

- A **marginal value function** $\boldsymbol{u_i}$ quantifies the preferences on **criterion** $\boldsymbol{g_i}$
  - *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

$$\underbrace{U(a)}_{\text{comprehensive value}} = \sum_{i=1,\ldots,n} \underbrace{w_i}_{\text{weight associated with criterion } g_i} \cdot \underbrace{u_i}_{\text{marginal value function associated with } g_i}(\underbrace{g_i(a)}_{\text{performance of alternative } a \text{ on } g_i}) = w_1 \cdot u_1(g_1(a)) + \ldots + w_n \cdot u_n(g_n(a))$$

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

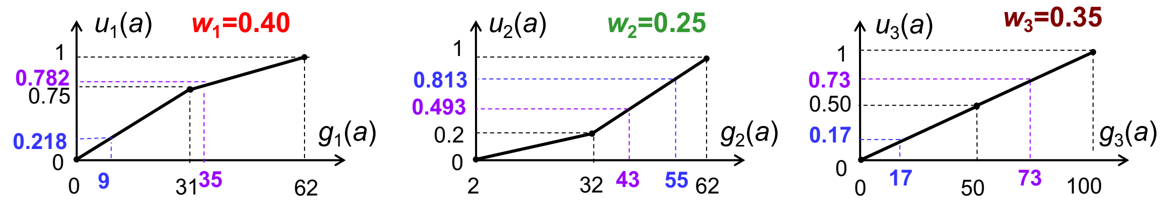
Value function function distinguishes only 2 possible relations between alternatives:

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

[5] Cząstkowe funkcje wartości przedstawiają preferencje na każdym kryterium, wykorzystując swego rodzaju skalę przedziałową pomiędzy 0 a 1. Aby ocenić warianty, musimy jednak przejść od poziomu pojedynczego kryterium do warstwy obejmującej wszystkie kryteria jednocześnie. W tym celu musimy zagregować cząstkowe wartości w jedną globalną. Na tym etapie stosujemy intuicyjną zasadę mówiącą, że ważniejsze kryteria powinny mieć większy wpływ na wartość globalną. Wartość ta jest definiowana jako suma ważona wartości cząstkowych dla wszystkich kryteriów. Model zbiera więc wkłady ze wszystkich kryteriów i sumuje je, jednocześnie różnicując ich wpływ. Należy pamiętać, że ma to sens tylko wtedy, gdy preferencje dotyczące każdego kryterium są niezależne od pozostałych kryteriów. Taki model oparty na wartościach jest w pełni kompensacyjny. Oznacza to, że dobre wartości na jednym kryterium mogą kompensować złe wartości na innych kryteriach. Ponadto, gdy każdemu wariantowi ostatecznie przypisujemy pojedynczą liczbę, mogą zaistnieć tylko relacje preferencji i nierozróżnialności. Preferencja zachodzi, gdy jeden wariant ma wyższą wartość globalną niż inny, natomiast nierozróżnialność wymaga takich samych wartości globalnych.

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*

$$U(A) = \sum_{i=1,\ldots,n} w_i \cdot u_i(g_i(A)) = w_1 \cdot u_1(26) + w_2 \cdot u_2(40) + w_3 \cdot u_3(44) =$$
$$= 0.40 \cdot 0.629 + 0.25 \cdot 0.413 + 0.35 \cdot 0.440 = 0.509$$

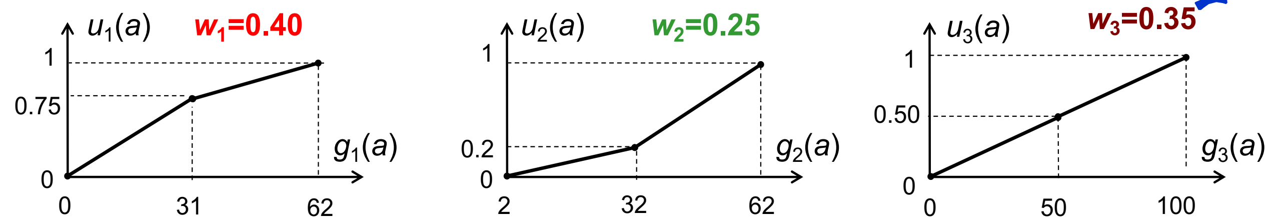
[6] Sprawdźmy, jak interpretować taką addytywną funkcję wartości. W górnej części slajdu znajdują się cząstkowe funkcje wartości dla trzech kryteriów. Ich wagi sugerują, że największy wpływ na globalną wartość ma pierwsze kryterium, a najmniejszy drugie. Aby uporządkować warianty od najlepszego do najgorszego, musimy odczytać cząstkową wartość dla każdego kryterium. Gdy oceny wariantów odpowiadają punktom charakterystycznym funkcji wartości, jest to łatwe. W przeciwnym razie musimy zastosować interpolację liniową. Na przykład, ocena 35 na pierwszym kryterium znajduje się pomiędzy 31, któremu przypisano wartość 0,75, a 62, któremu przypisano wartość 1. Ponieważ ocena 35 znajduje się w czterech trzydziestych pierwszych podzakresu, przypisujemy jej wartość 0.782. W drugim etapie cząstkowe wartości są agregowane do wartości globalnych przy użyciu modelu sumy ważonej. Przykład takiej agregacji możecie zobaczyć na dole slajdu dla kandydata Andreev, wariant A. Na koniec wszyscy kandydaci są porządkowani w kolejności malejących globalnych wartości. W naszym przypadku wybralibyśmy Dubova, wariant D, o najwyższej wartości globalnej, a na pewno odrzucilibyśmy Jope’a, wariant J, który znajduje się na samym dnie rankingu.

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:

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

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

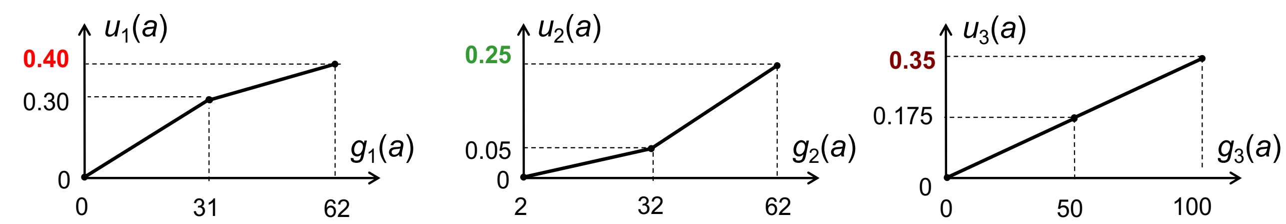


It is possible to incorporate weights into marginal value function by assuming $w_i = u_i(\beta_i)$, for all criteria

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

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

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



[7] Dotychczas rozważaliśmy model, w którym każda funkcja cząstkowa przyjmowała wartości od 0 do 1, a wagi kontrolowały wpływ danego kryterium na wynik całkowity. Takie wagi można interpretować jako przetargi między wartościami cząstkowymi. Możliwe jest jednak włączenie wag do funkcji cząstkowych poprzez założenie, że maksymalna wartość cząstkowa odpowiadająca najlepszej ocenie na danym kryterium jest równa wadze. Pozostałe wartości są wtedy odpowiednio skalowane poprzez pomnożenie ich przez wagi. Widać to wyraźnie dla naszego problemu. Tam, gdzie na pierwszym kryterium mieliśmy wartość 1, mamy teraz 0.4; tam, gdzie mieliśmy 0.75, mamy teraz 0.3. Mając tak przeskalowane funkcje cząstkowe, możemy obliczyć wartość globalną, sumując wartości cząstkowe, co jest jeszcze łatwiejsze w interpretacji niż w poprzednio przyjętym modelu.

Need for Easy Preference Information

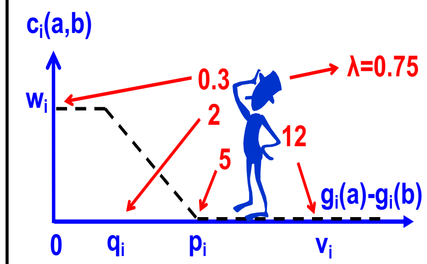
Traditional MCDA methods require **rich (and sometimes) difficult** preference information:

- 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

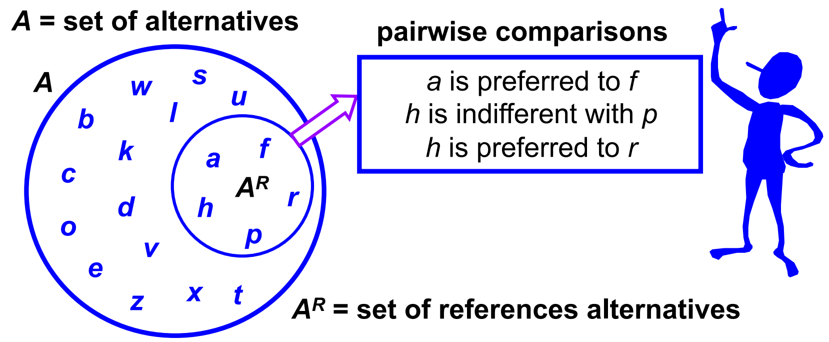


- The traditional MCDA methods may be **too demanding of cognitive effort** of their users
- We advocate for methods requiring „easy" preference information
- **"Easy" means natural and even partial**

[8] Wiemy już, jak stosować addytywną funkcję wartości, ale nie wiemy, jak ją utworzyć. Tradycyjne metody wymagają bezpośredniego określenia wartości wiele parametrów. Aby to zrobić, musimy rozumieć ich znaczenie i nawet jeśli większość z nich jest wysoce interpretowalna, w niektórych przypadkach możemy być przeciążeni dużą liczbą parametrów. W związku z tym, istotny nurt we wspomaganiu decyzji polega na rozwijaniu podejść wymagających łatwych preferencji, przy czym łatwe oznacza naturalne i częściowe.

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

[9] W trakcie tego wykładu będziemy wykorzystywać pośrednią informację preferencyjną w formie holistycznych ocen wariantów. Istnieją badania psychologiczne potwierdzające, że decydenci są bardziej pewni gdy przedstawiają takie decyzje, ponieważ jest to coś, do czego są przyzwyczajeni. W kontekście rankingu i wyboru, najbardziej naturalną formą pośrednich preferencji są holistyczne porównania parami. Wskazują one na preferencję jednego wariantu nad innym lub nierozróżnialność pary wariantów. Takie oceny dotyczą podzbioru tzw. wariantów referencyjnych. Mogą to być warianty stosunkowo dobrze znane decydentowi, warianty, dla których już w przeszłości wyrażał on preferencje, lub opcje fikcyjne, specjalnie utworzone, aby łatwiej było je holistycznie porównać. W każdym z tych przypadków decydent jest proszony o przedstawienie lub potwierdzenie swoich holistycznych preferencji, biorąc pod uwagę oceny wariantów referencyjnych na wszystkich kryteriach.

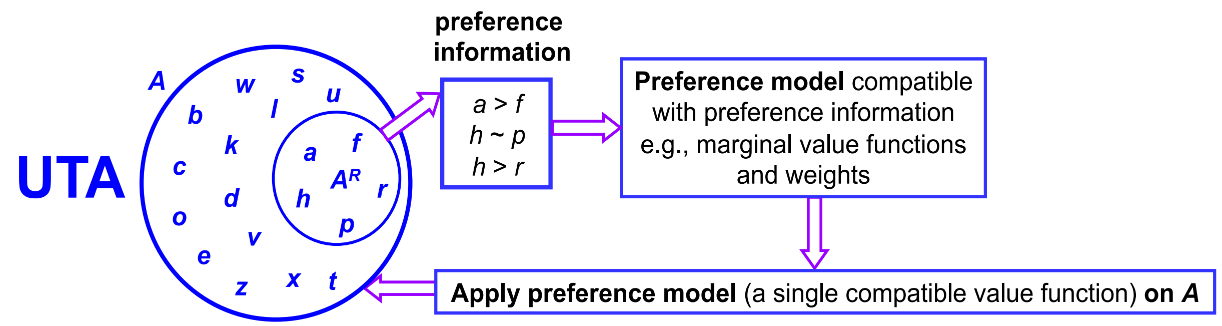
The UTA Method

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

[10] Metody takie jak PROMETHEE są oparte na tradycyjnym paradygmacie agregacji. Zakłada on, że model preferencji jest najpierw konstruowany na podstawie bezpośrednich preferencji, a następnie stosowany do zbioru wariantów w celu wypracowania rekomendacji. Dziś skupimy się na paradygmacie dezagregacji-agregacji. Zgodnie z jego założeniami znane są holistyczne preferencje decydenta dla podzbioru wariantów referencyjnych i służą one do konstrukcji spójnego z nimi modelu preferencji, to jest takiego, który pozwala na ich odtworzenie. Dopiero potem taki model jest stosowany do oceny wszystkich wariantów. Inną nazwą tego nurtu metod jest regresja porządkowa. W pewnym sensie ma ona na celu odkrycie intencji decydenta przy dokonywaniu przez niego holistycznych ocen. Jest to zgodne z postulowaną przez Marcha tzw. zasadą racjonalności ex post oraz, co ważniejsze, z paradygmatem uczenia się z przykładów. Ten ostatni leży u podstaw współczesnej sztucznej inteligencji z jej dwoma kluczowymi dyscyplinami: uczeniem maszynowym i odkrywaniem wiedzy. Następne kilkanaście slajdów spędzimy na omówieniu lub przypomnieniu metody UTA, która implementuje paradygmat dezagregacji preferencji.

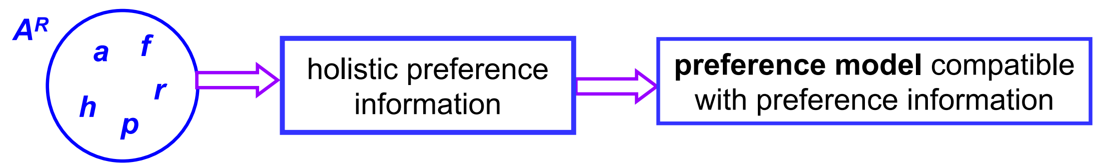
Preference Disaggregation

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)



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*

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

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

[11] UTA wymaga pośredniej informacji preferencyjnej. W oryginalnej postaci zaproponowanej w latach 80-tych metoda ta tolerowała zupełny ranking wariantów referencyjnych, które musiały być uporządkowane od najlepszego do najgorszego. Jednak łatwo jest uogólnić to podejście, aby dopuścić częściowy ranking w postaci porównań parami. Informacja taka jest wykorzystywana do konstrukcji spójnego modelu preferencji, którym w przypadku UTY jest addytywna funkcja wartości z monotonicznymi cząstkowymi funkcjami na poszczególnych kryteriach. Konstrukcja ta opiera się na obserwacji, że gdy decydent stwierdza, że wariant a jest preferowany nad wariant b, jego globalna użyteczność powinna być większa; natomiast w przypadku, gdy dwa warianty są oceniane jako nierozróżnialne, ich globalne wartości powinny być równe.

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:

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

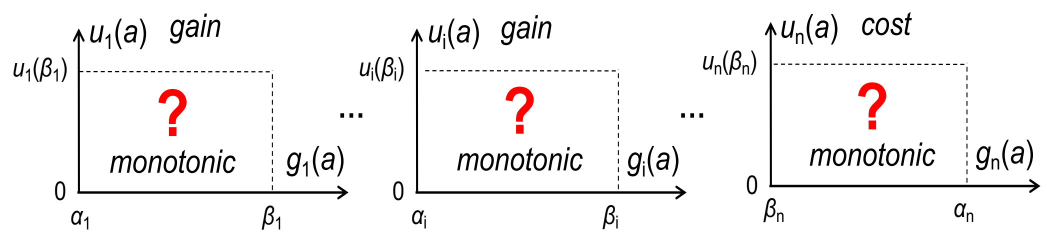
- For gain-type criteria, α_i is the least performance and for cost-type criteria α_i is the greatest performance;

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

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

- 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



$$\boxed{u_i(\alpha_i) = 0 \quad \forall i \in I \qquad \sum_{i=1}^{n} u_i(\beta_i) = 1}$$

[12] Proces konstrukcji takiej funkcji podlega ograniczeniom, które zapewniają wysoką interpretowalność modelu. Ich pierwszy rodzaj związany jest z normalizacją wartości globalnych do przedziału między 0 a 1. Można to łatwo osiągnąć, przyjmując, że wartość cząstkowa dla najgorszej oceny na każdym kryterium musi wynosić zero, a suma wartości cząstkowych odpowiadających najlepszym ocenom na wszystkich kryteriach wynosi jeden. W ten sposób wariant anty-idealny uzyskuje globalną wartość 0, a wariant idealny otrzymuje wartość 1. Podkreślmy, że te skrajne oceny można określić w dwojaki sposób. Z jednej strony mogą to być najgorsze i najlepsze oceny zaobserwowane w zbiorze wariantów, co jest założeniem w naszym przykładzie. Z drugiej strony mogą to być oceny ograniczające skalę, na której można wyrazić osiągi na danym kryterium. Na przykład, można by użyć skali od 0 do 100, a takie skrajne oceny nie muszą być osiągnięte przez żaden rozważany wariant. Możemy jednak chcieć, aby funkcja cząstkowa była zdefiniowana w całym przedziale zmienności ocen.

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

Reconstruction works for any:

$u_1(10) > u_2(10)$

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

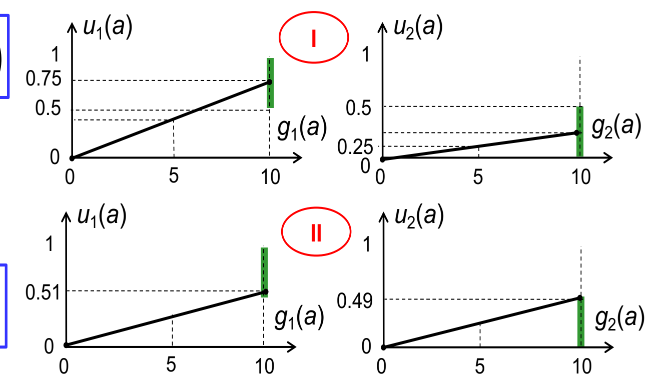
$U(X) = \frac{10}{10} u_1(10) + \frac{0}{10} u_2(10) = u_1(10)$

v

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

v

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



	$u_1(a)$	$u_2(a)$	$U(a)$
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

[13] Aby zrozumieć to, co omówiliśmy do tej pory, rozważmy prosty problem z dwoma kryteriami typu zysk ze skalą od 0 do 10, trzema wariantami X, Y i Z oraz liniowymi funkcjami wartości. Ranking referencyjny, który chcemy odtworzyć, to: X jest preferowany nad Y, a Y jest preferowany nad Z. W przypadku wykorzystania funkcji liniowych możliwe jest modelowanie każdej cząstkowej wartości jako odpowiedniej proporcji maksymalnej cząstkowej wartości odpowiadającej najlepszej ocenie na danym kryterium. Analizując oceny X, Y i Z, możemy stwierdzić, że ranking referencyjny zostanie odtworzony przy funkcjach cząstkowych zapewniających, że maksymalna wartość na pierwszym kryterium jest większa niż maksymalna wartość na drugim kryterium. Pamiętamy jednak, że muszą one zostać znormalizowane, aby sumowały się do 1. Na tym slajdzie najpierw przedstawiłem funkcje cząstkowe takie, że maksymalny udział dla pierwszego kryterium wynosi 0.75, a dla drugiego – 0.25. Widać, że ranking referencyjny został odtworzony, ponieważ X osiąga wartość globalną większą niż Y, a Y większą wartość niż Z. Ten sam efekt obserwujemy dla funkcji cząstkowych o maksymalnych wartościach 0.51 i 0.49. Różnice między wartościami globalnymi wariantów X, Y i Z są tu jednak mniejsze.

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*

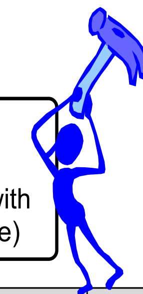
Alt.	$g_1 \uparrow$	$g_2 \uparrow$
X	10	0
Y	5	5
Z	0	10

• *X* > • *Y* > • *Z* ⇨ • *X* > • *Z* > • *Y*

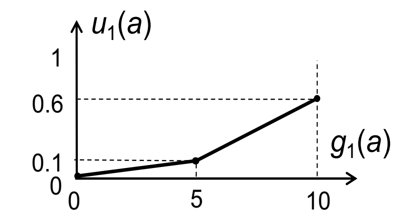
It is not possible to reconstruct the ranking with linear marginal value functions, i.e., to find $u_1(10)$ and $u_2(10)$ satisfying the constrains

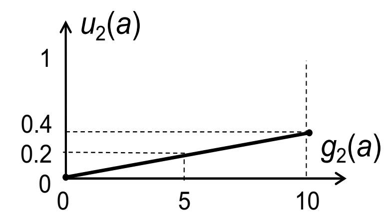
$$\boxed{U(X) = u_1(10)} > \boxed{U(Z) = u_2(10)} > \boxed{U(Y) = 0.5u_1(10) + 0.5u_2(10)}$$

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



- 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

[14] Rozważmy teraz ten sam problem, ale z nieco zmienionym rankingiem referencyjnym, w którym wariant Z powinien być uplasowany wyżej niż Y. Rozważenie tych preferencji prowadzi do wniosku, że nie da się ich odtworzyć za pomocą liniowych funkcji wartości, ponieważ preferencje X nad Z oraz Z nad Y są sprzeczne. Dlatego liniowy charakter funkcji cząstkowych może w niektórych przypadkach nie być wystarczający. Wtedy rozwiązaniem powinno być zwiększenie elastyczności funkcji poprzez dodanie tzw. punktów charakterystycznych. Może w nich zmieniać się nachylenie funkcji. Dla rozważanego problemu wystarczyłoby dodać jeden punkt w środku pierwszego kryterium, co umożliwiłoby przypisanie nieco mniejszej wartości cząstkowej do oceny 5. Wtedy odtworzenie rankingu referencyjnego: X > Z > Y staje się możliwe.

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

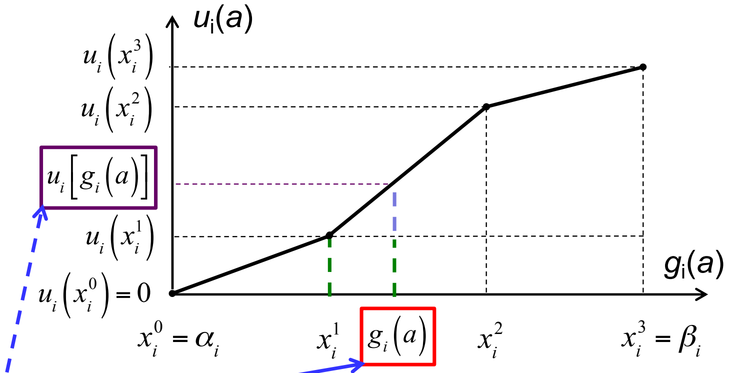
$$x_i^j = \alpha_i + \frac{j-1}{\gamma_i - 1}(\beta_i - \alpha_i), \;\; j = 0, \dots, \gamma_i$$

- 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**:

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

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



- 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

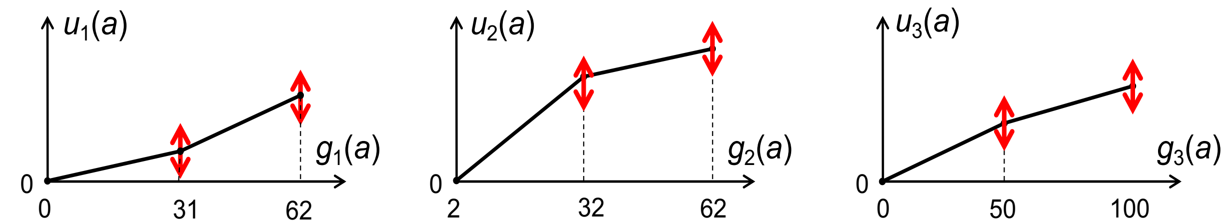
[15] Metoda UTA wykorzystuje cząstkowe funkcje wartości, które są odcinkami liniowe, przy czym liczba tych odcinków jest podana przez decydenta. Jest ona określona przez liczbę tzw. punktów charakterystycznych. Taka liczba jest elementem informacji preferencyjnej. Dla funkcji liniowej mamy dwa takie punkty; dla 2-odninkowej - 3 punkty, itd. Domyślnie UTA przyjmuje, że punkty te są równomiernie rozłożone, dzieląc zakres oceny na równe podprzedziały. Założenie stosowania funkcji odcinkami liniowych implikuje sposób obliczania wartości cząstkowych. Gdy oceny odpowiadają punktom charakterystycznym, łatwo jest odczytać odpowiadające im wartości. Gdy ocena znajduje się pomiędzy dwoma punktami charakterystycznymi, wartość oblicza się za pomocą interpolacji liniowej. Zaczynamy od wartości cząstkowej gorszego sąsiedniego punktu charakterystycznego, a następnie zwiększamy ją o odpowiednią część różnicy między wartościami cząstkowymi przypisanymi dwóm sąsiednim punktom. Co oznacza "odpowiednią" wynika bezpośrednio z twierdzenia Talesa.

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

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

[16] UTA modyfikuje wartości cząstkowe przypisane punktom charakterystycznym, aby dopasować je do informacji preferencyjnej. Są to zmienne rozpatrywane w problemie regresji porządkowej. Przy jego rozwiązywaniu UTA zapewnia interpretowalność modelu. Z jednej strony, wartości tych zmiennych muszą być nieujemne. Z drugiej muszą one respektować monotoniczność cząstkowych funkcji wartości. To ostatnie zapewnia się poprzez rozważenie wszystkich sąsiadujących punktów charakterystycznych. Wartość przypisana do lepszego punktu musi być co najmniej tak dobra jak wartość przypisana do punktu gorszego. Oczywiście to, co znaczy lepszy lub gorszy, zależy od tego, czy rozpatrujemy kryterium typu zysk lub koszt. Funkcje cząstkowe muszą być niemalejące dla kryteriów zysk i nierosnące dla kryteriów koszt. Niezależnie od typu, przebieg funkcji cząstkowej jest nieściśle monotoniczny.

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

$$\boxed{\begin{array}{l} U(a) \geq U(b) + \varepsilon \Leftrightarrow a > b \\ U(a) = U(b) \quad \Leftrightarrow a \sim b \end{array}}$$

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

[17] Wiemy już, że problem regresji porządkowej rozwiązywany przez UTA musi gwarantować normalizację i monotoniczność addytywnej funkcji wartości. Przede wszystkim musimy jednak odtworzyć porównania parami dostarczane przez decydenta. Uświadomiłem już wam, że takie odtworzenie nie zawsze musi być możliwe. Dlatego powinniśmy dać dezagregacji preferencji szansę na popełnienie pewnego błędu. Na dzisiejszym wykładzie przedstawię uproszczony sposób, który różni się od oryginalnej propozycji w metodzie UTA. Założymy, że nierozróżnialności wariantów referencyjnych muszą być odtworzone, a przy modelowaniu preferencji nierówność ostrą zmienimy w nieostrą, dodając po jej prawej stronie epsilon. Moglibyśmy przyjąć, że jest ta pewna bardzo mała stała dodatnia i wszystko byłoby w porządku. My potraktujemy jednak tego epsilona jako zmienną tak, by sprawdzić czy istnieje funkcja wartości spójna z preferencjami decydenta.

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**

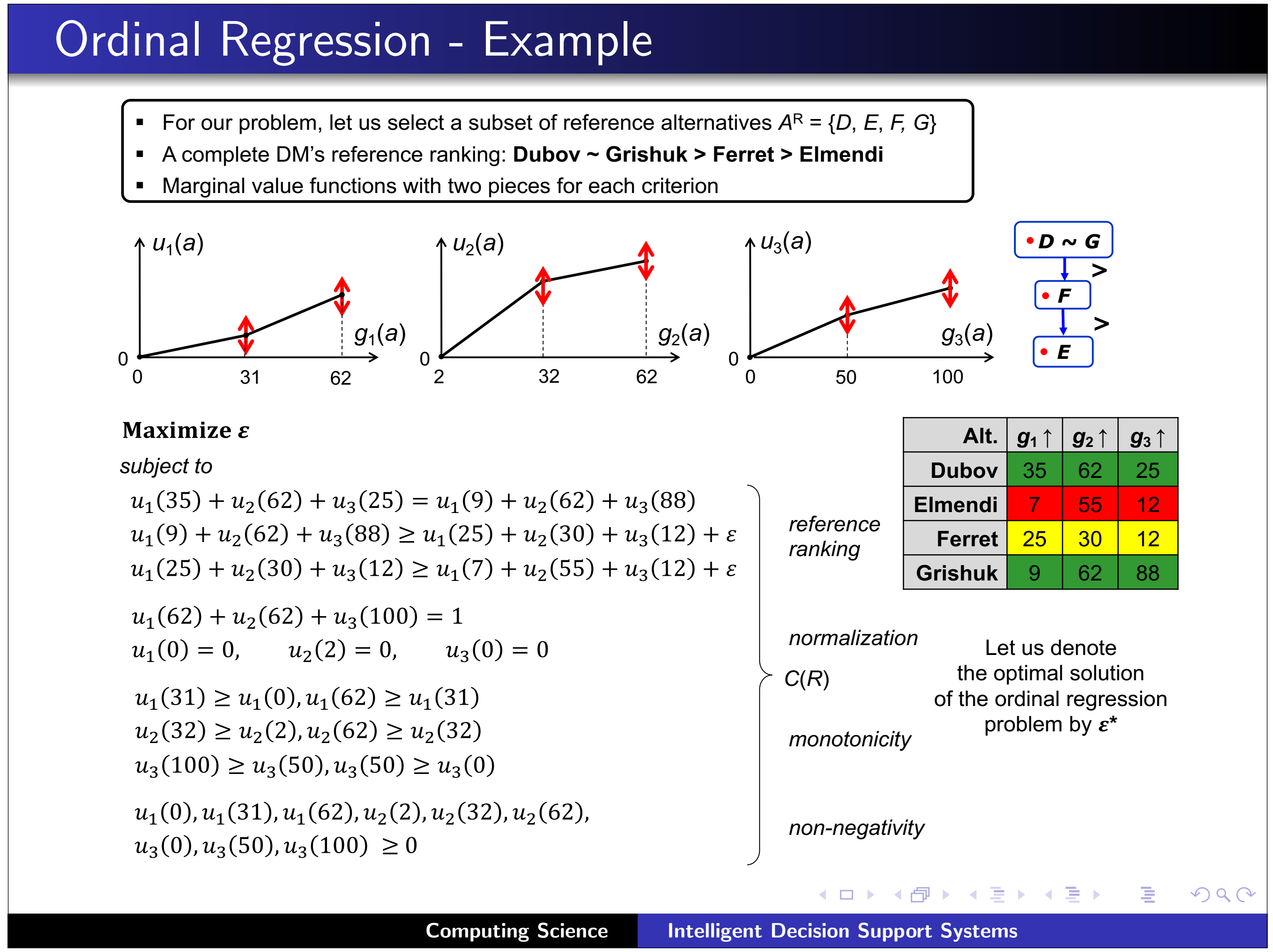
subject to

$$
\left.\begin{array}{l}
\left.\begin{array}{ll}
U(a) \geq U(b) + \varepsilon \Leftrightarrow a \succ b & \forall a,b \in A^R \\
U(c) = U(d) \quad \Leftrightarrow c \sim d & \forall c,d \in A^R
\end{array}\right\} R \\
\left.\begin{array}{ll}
\sum_{i=1}^{n} u_i(\beta_i) = 1 & \\
u_i(\alpha_i) = 0, \quad \forall i \in I & \\
u_i(x_i^{j+1}) - u_i(x_i^{j}) \geq 0, & j = 0, \dots, \gamma_i - 1;\ \forall i \in I \\
u_i(x_i^{j}) \geq 0,\ \forall\, i \text{ and } j &
\end{array}\right\} C
\end{array}\right\} C(R)
$$

reference ranking (first two constraints); *normalization* ($\sum_{i=1}^{n} u_i(\beta_i)=1$, $u_i(\alpha_i)=0$); *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.

[18] Zbierzmy więc wszystkie omówione dotychczas aspekty w jednym miejscu i sformułujmy problem regresji porządkowej, który jest silnikiem metody UTA. Jest to model programowania liniowego. Skupmy się najpierw na ograniczeniach. Chcemy znaleźć addytywną funkcję wartości, która: odtwarza ranking referencyjny, jest znormalizowana do przedziału od 0 do 1 oraz stosuje monotoniczne cząstkowe funkcje wartości. Tutaj przedstawiam ograniczenia tylko dla kryteriów typu zysk. W ograniczeniach tłumaczących ranking referencyjny, dla relacji preferencji wykorzystujemy epsilona, a relacje nierozróżnialności są tłumaczone jako równości odpowiednich użyteczności globalnych. W funkcji celu maksymalizujemy wartość zmiennej epsilon. Oznacza to, że chcemy znaleźć model, który jest znormalizowany i monotoniczny, ale też odtwarza wszystkie porównania parami. Stanie się tak, gdy optymalna wartość epsilona będzie większa od zera. Wtedy bowiem będzie zapewniał ostry, ścisły charakter nierówności dla wszystkich relacji preferencji.



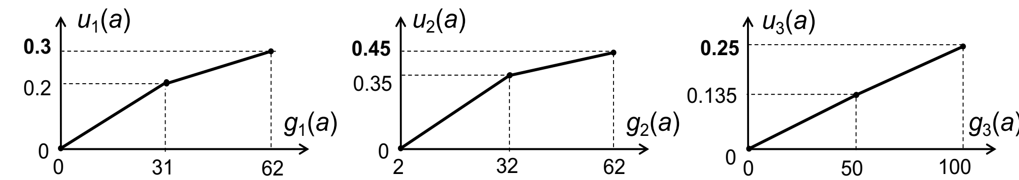
[19] Ogólne sformułowanie problemu jest świetne, ale przykład jest jeszcze lepszy. Wróćmy do naszego problemu i wybierzmy Dubova, Elmendiego, Ferret i Grishuka jako nasz zbiór referencyjny. Jako decydenci oceniamy Dubova i Grishuka najlepiej, lepiej od Ferret, a tego z kolei lepiej od Elmendiego. Zakładamy też, że cząstkowe funkcje wartości na wszystkich kryteriach korzystają z trzech punktów charakterystycznych; to znaczy, że można je złamać raz, dokładnie w środku. W problemie regresji porządkowej modelujemy relację nierozróżnialności dla pary (Dubov, Grishuk) oraz relacje preferencji dla dwóch par sąsiadujących w rankingu referencyjnym: (Grishuk, Ferret) oraz (Ferret, Elmendi). Każda globalna wartość jest modelowana jako suma cząstkowych wartości odpowiadających ocenom poszczególnych kandydatów. Po prawdzie dla ocen, które nie odpowiadają punktom charakterystycznym powinniśmy też zamodelować odpowiadające im cząstkowe wartości w funkcji wartości przypisanych sąsiednim punktom charakterystycznym. Dalej zapewniamy spełnienie ograniczeń normalizacji i monotoniczności. Wszystkie trzy kryteria są typu zysk. Na przykład, dla kryterium drugiego, ocenie dwa przypisujemy wartość zerową, a wartość przypisana ocenie 32 musi być mniejsza lub równa od wartości przypisanej ocenie 62. Po rozwiązaniu takiego problemu otrzymujemy maksymalną wartość zmiennej epsilon. Nazwijmy ją epsilon z *.

Compatible Value Function (1)

- **If ε^*>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

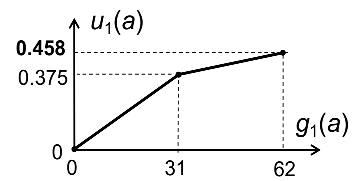
[20] Jeśli ta optymalna wartość jest większa od zera, to przynajmniej jedna funkcja wartości jest spójna z dostarczoną informacją preferencyjną. Na tym slajdzie przedstawiłem jedną instancję takiego modelu. Składa się on z funkcji cząstkowych wyznaczonych przez wartości przypisane trzem punktom charakterystycznym na każdym kryterium. Skupmy się najpierw na tych funkcjach. Są one monotoniczne oraz respektują charakterystykę typu zysk. Są one również znormalizowane. Najgorszym ocenom przypisano wartość cząstkową równą zero, natomiast najlepszym ocenom przypisano wartości 0.3, 0.45 i 0.25, które sumują się do 1. Dla wszystkich dwunastu wariantów, od Andreeva do Liu, od A do L, jesteśmy w stanie odczytać wartości cząstkowe. Na przykład dla Andreeva sumują się one do wartości globalnej nieco mniejszej niż 0.7. Należy pamiętać, że teraz wagi są niejawne i to maksymalne udziały funkcji cząstkowych kontrolują wpływ kryteriów na ostateczny wynik. Wartości pozostałych wariantów mogą być obliczone analogicznie. W szczególności Dubov i Grishuk otrzymują globalną wartość równą 0.75, Ferret – 0.520, a Elmendi – wartość 0.504. Wszystkie one przyczyniają się do odtworzenia rankingu referencyjnego. Efektem ubocznym wykorzystania tego modelu jest jednak to, że pozwala on na ocenę wszystkich wariantów, także tych niereferencyjnych, które nie były bezpośrednio ocenione przez decydenta. Pozycje wszystkich dwunastu kandydatów podane są w tabeli. W tym przypadku Dubov i Grishuk znajdują się na szczycie rankingu.

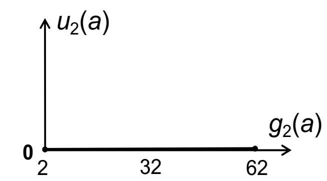
Compatible Value Function (2)

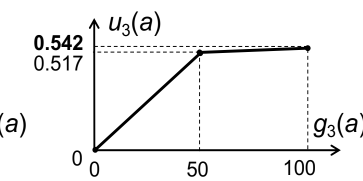
- Let us inspect another solution (e.g., corresponding to max ε^*)
- Increased maximal share in the comprehensive value of criterion g_1 and g_3 (the shapes are more concave) and zeroed maximal share of criteria g_2

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:

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

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

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

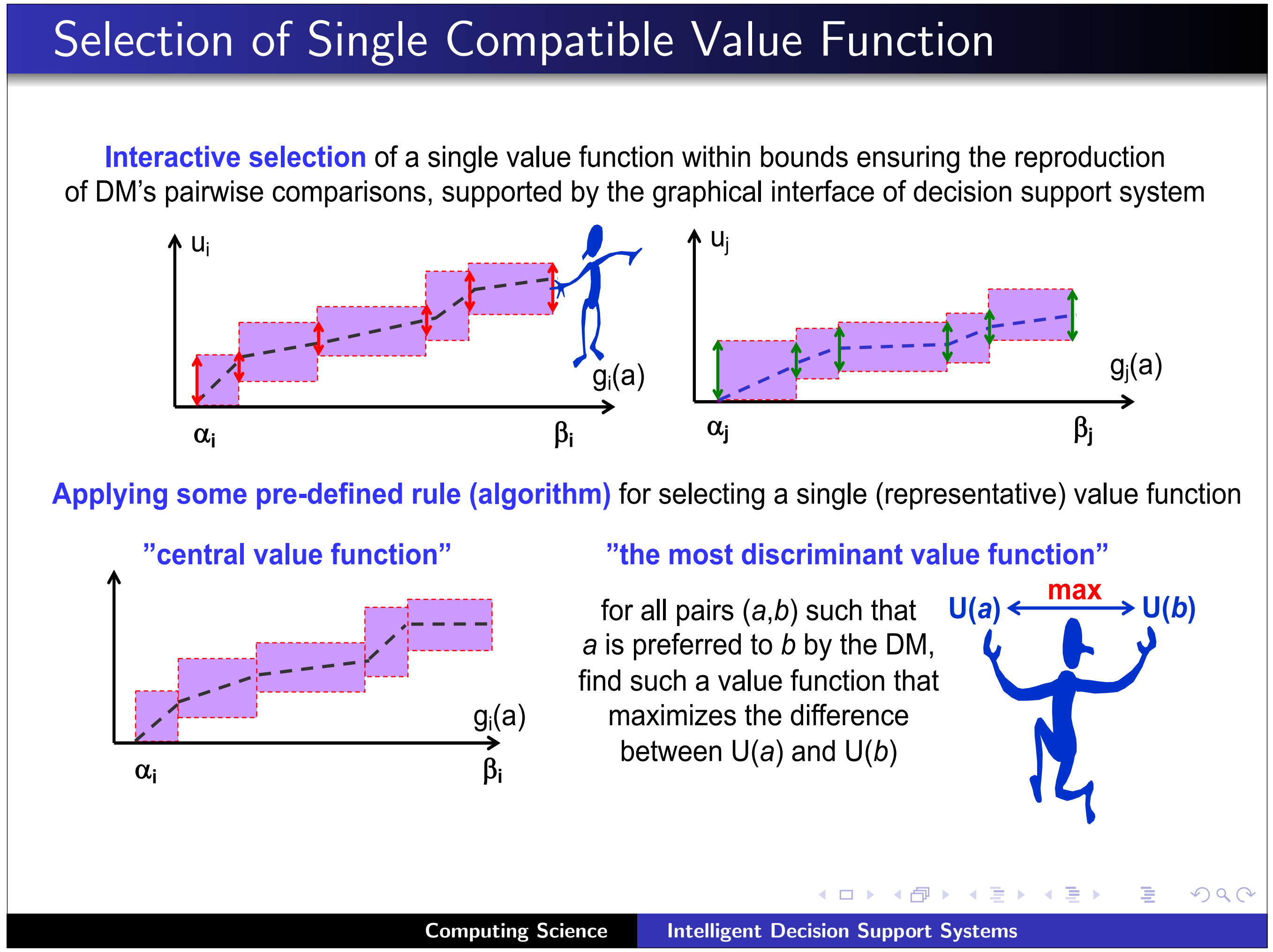
$U(E) = u_1(7) + u_2(55) + u_3(12) = 0.209$

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

[21] Gdy zbiór możliwych rozwiązań dla problemu regresji porządkowej jest niepusty, to zwykle istnieje więcej niż jedna, a nawet nieskończenie wiele funkcji wartości spójnych z preferencjami decydenta. Na tym slajdzie przedstawiam inną taką funkcję, odpowiadającą maksymalnej wartości zmiennej epsilon. Ponownie jest ona monotoniczna i znormalizowana. Różni się jednak w stosunku do poprzedniej funkcji choćby zwiększonym maksymalnym udziałem pierwszego i trzeciego kryterium oraz wyzerowanym wkładem kryterium drugiego. Nawet jeśli wartości cząstkowe i globalne czterech wariantów referencyjnych są tu inne, to ranking referencyjny jest nadal odtworzony. Wybór spójnej funkcji może jednak wpłynąć na ranking wszystkich wariantów. I rzeczywiście, teraz to Andreev jest najwyżej sklasyfikowanym kandydatem, a Brown znajduje się w rankingu wyżej niż Calvet. Tych różnic jest jednak zdecydowanie więcej.



[22] Gdy istnieje wiele rozwiązań problemu regresji porządkowej, powstaje pytanie, jak sobie z taką mnogością poradzić. Istnieją trzy podstawowe odpowiedzi. Po pierwsze, wybór można pozostawić użytkownikowi. Istnieją systemy wspomagania decyzji z graficznym interfejsem pozwalające na interaktywne dopasowanie funkcji wartości do własnych potrzeb w ramach dopuszczalnych granic. Po drugie, możemy skorzystać z pewnych predefiniowanych reguł wyboru jednej, spójnej funkcji wartości. W literaturze można znaleźć dziesiątki sensownych algorytmów służących do tego celu. Wśród nich dwa można łatwo wyjaśnić. Z jednej strony sensowny jest wybór modelu centralnego o średnich wartościach cząstkowych obliczonych na podstawie wartości skrajnych możliwych do uzyskania modeli. Z drugiej strony, sensowne jest też maksymalne podkreślenie konsekwencji porównań parami dokonywanych przez decydenta. To znaczy, jeśli twierdził on, że a jest preferowane nad b, to zgodnie z modelem przewaga a nad b powinna być jak największa. To odpowiada właśnie maksymalizacji epsilona. Trzecie rozwiązanie polega na uwzględnieniu wszystkich spójnych funkcji wartości jednocześnie. I to właśnie na nim skupimy się w dalszej części wykładu.

The UTA Method – Summary

Preference information UTA

- 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

Preference model:

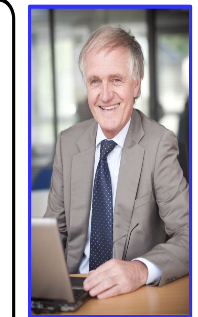
- Additive value function with piecewise linear marginal value functions

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

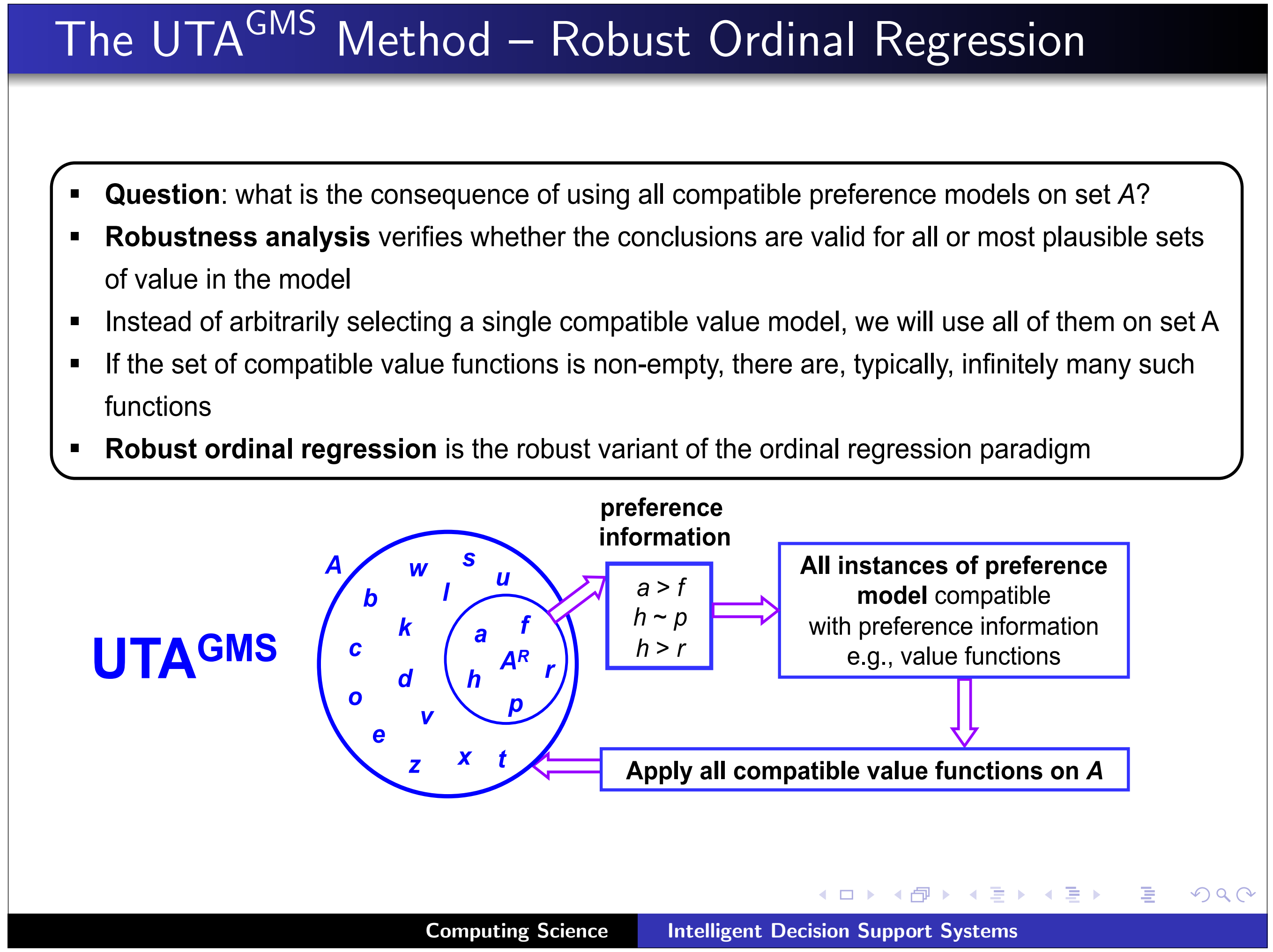
Recommendation:

- 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.

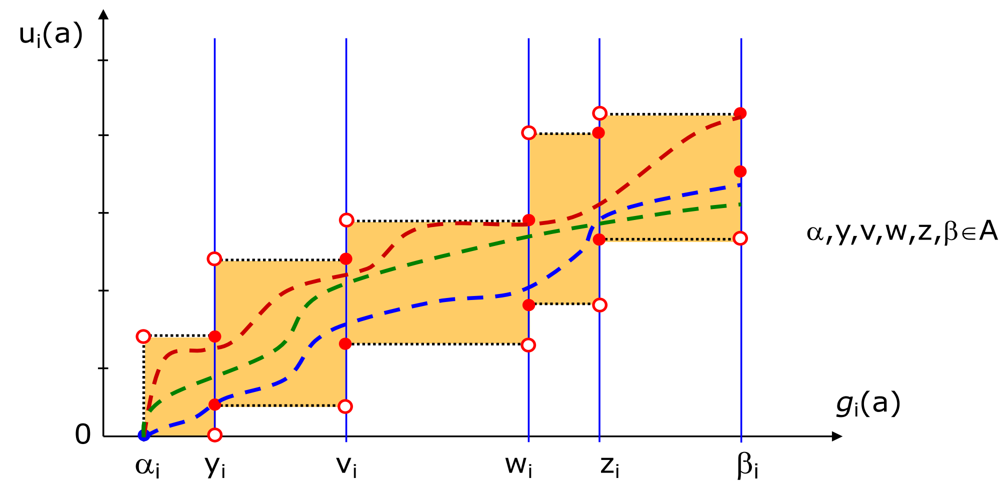
[23] Zanim do tego dojdzie, podsumujmy własności metody UTA. Akceptuje ona pośrednią informację preferencyjną w postaci zupełnego rankingu dla podzbioru wariantów referencyjnych. Można ją jednak łatwo dostosować do wykorzystania porównań parami. Takie holistyczne oceny są tłumaczone do spójnego z nimi modelu preferencji. UTA stosuje addytywną funkcję wartości z funkcjami cząstkowymi o charakterze odcinkami liniowym. Uczenie modelu odbywa się przez rozwiązanie problemu programowania liniowego zwanego regresją porządkową. Wyprowadzony model może być wykorzystany do oceny wszystkich wariantów i uporządkowania ich od najlepszego do najgorszego.



[24] Przejdźmy teraz do umówienia rozszerzenia metody UTA o nazwie UTA-GMS. Wykorzystuje ona porównania parami wariantów referencyjnych i stara się odpowiedzieć na pytanie o konsekwencje zastosowania wszystkich funkcji spójnych z takimi preferencjami do uszeregowania zbioru wariantów. W tym celu przeprowadza analizę odporności, a więc badanie stabilności oferowanego wyniku w zależności od wyboru spójnej instancji założonego modelu preferencji. Unikamy więc arbitralnego wyboru jednej funkcji, zamiast tego skupiając się na jednoczesnym wykorzystaniu ich wszystkich, a takich funkcji może być nieskończenie wiele. Ten sposób postępowania nazywa się odporną regresją porządkową.

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**

[25] Jak zdefiniować zbiór wszystkich funkcji spójnych z preferencjami decydenta? Wybór funkcji liniowych lub odcinkami liniowych jest w pewnym sensie arbitralny. Aby uwzględnić wszystkie możliwe funkcje, trzeba ich postać uogólnić tak, by punkty charakterystyczne odpowiadały wszystkim unikalnym ocenom wariantów. Wartości cząstkowe w tych punktach są nieznane, ale porównania wariantów parami narzucają na nie ograniczenia tak, że w każdym punkcie dozwolone wartości są pewnym przedziałem. Razem przedziały te wyznaczają obszar, w którym mogą przebiegać spójne funkcje wartości. Trzeba tylko zapewnić, by funkcje te były monotoniczne.

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

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

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

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

$$u_i(x_i^{j+1}) - u_i(x_i^j) \geq 0, j = 1, \dots, n_i(A) - 1; \ \forall i \in I \quad \text{monotonicity}$$

$$u_i(\beta_i) \geq u_i(x_i^j) \geq u_i(\alpha_i), \ \forall\ i \text{ and } j \quad \text{bounds (non-negativity)}$$

(constraints from normalization to bounds: C; all constraints: $C(R)$)

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

[26] Pierwszy etap działania metody polega na sprawdzeniu spójności informacji preferencyjnej z założonym modelem preferencji. Rozwiązywany problem jest bardzo podobny do tego, który omówiliśmy dla nieco zmodyfikowanej metody UTA. Dwie główne różnice polegają na tym, że nie tłumaczymy już zupełnego rankingu referencyjnego, a porównania parami, a punkty charakterystyczne funkcji cząstkowych o charakterze ogólnym są zdefiniowane we wszystkich unikalnych ocenach wariantów na poszczególnych kryteriach. Jeśli optymalna wartości epsilona jest większa od zera, to istnieje co najmniej jedna funkcja spójna z preferencjami decydenta. Jeśli warunek ten jest spełniony, to zazwyczaj istnieje wiele, a nawet nieskończenie wiele, takich funkcji. Pytanie, jak można sprawdzić, czy prowadzą one do spójnych rekomendacji.

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* ($a \succeq^N b$) iff $U(a) \geq U(b)$ **for all** compatible value functions
- $\succeq^N$ 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
- ***a* is possibly preferred to *b*** ($a \succeq^P b$) iff $U(a) \geq U(b)$ **for at least one** compatible value function
- $\succeq^P$ 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

N

P

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

[27] Podstawowy pomysł opiera się na konstrukcji dwóch relacji preferencji, koniecznej i możliwej. Relacja konieczna zachodzi dla pary wariantów (a,b), jeżeli wariant a jest co najmniej tak samo dobry jak wariant b dla wszystkich spójnych funkcji wartości. Oznacza to, że ma globalną wartość zawsze większą równą. To z kolei oznacza, że w zbiorze spójnych funkcji wartości minimalna różnica między wartościami globalnymi wariantów a oraz b jest większa równa od zera oraz, że nie jest istnieje żadna funkcja, dla której wariant b byłby ściśle lepszy od wariantu a. Relacja możliwa zachodzi dla pary wariantów (a,b), jeżeli a jest co najmniej tak dobry jak b dla co najmniej jednej spójnej funkcji wartości. Oznacza to, że jego globalna wartość jest większa równa dla co najmniej jednego dozwolonego scenariusza. To z kolei oznacza, że w zbiorze spójnych funkcji wartości maksymalna różnica między wartościami globalnymi wariantów a oraz b jest większa równa od zera. Zwróćcie uwagę, że te relacje mają charakter słaby, co oznacza, że może zachodzić relacja konieczna dla pary (a, b), a jednocześnie relacja możliwa dla pary (b, a). Wystarczy by dla co najmniej jednej funkcji, wartości tych wariantów były równe. Nie jest więc możliwe określenie jednej relacji na podstawie drugiej.

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 $\mathbf{D(a,b) = Max\ [U(a)-U(b)]}$ subject to $C(R)$ (ε set to a small positive value)
  $\mathbf{D(a,b) \geq 0}$ 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 $\mathbf{d(a,b) = Min\ [U(a)-U(b)]}$ subject to $C(R)$ (ε set to a small positive value)
  $\mathbf{d(a,b) \geq 0}$ 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

[28] Omówmy teraz sposoby weryfikacji prawdziwości tych dwóch relacji dla danej pary wariantów (a, b). Jeśli chodzi o relację możliwą, to można maksymalizować różnicę między globalnymi wartościami wariantów a oraz b przy zbiorze ograniczeń definiujących zbiór spójnych funkcji wartości. Jeśli taka maksymalna różnica jest większa równa od zera, to relacja możliwa jest prawdziwa. Wadą tego sposobu jest konieczność przypisania arbitralnej małej wartości do zmiennej epsilon tak, by zmieniła się ona w stałą. Drugi sposób unika tego. Opiera się on na dodaniu pewnej hipotezy do zbioru spójnych funkcji wartości, hipotezy, która mówi, że jest a może być co najmniej tak dobre jak b. Jeśli okaże się, że tak powstały zbiór ograniczeń jest niesprzeczny, a maksymalna wartości epsilona pozostanie większa od zera, oznacza to, że istnieje co najmniej jedna spójna funkcja wartości, dla której a jest co najmniej tak dobre jak b, a zatem zachodzi relacja możliwa. Dla relacji koniecznej sposób postępowania jest podobny. Zamiast maksymalizacji różnicy wartości globalnych, mamy minimalizację. Z kolei metoda z hipotezą musi być sprytniejsza. Hipoteza, którą stawiamy jest przeciwna do relacji, którą chcemy udowodnić. W zbiorze spójnych funkcji wartości zakładamy więc, że b może być ściśle lepsze od a. Jeśli wykażemy, że nie jest to możliwe, to a będzie koniecznie preferowane nad b. A to stanie się, gdy zbiór ograniczeń stanie się sprzeczny lub maksymalna wartość epsilona nie będzie większa od zera.

Properties of the Necessary and Possible Relations

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)

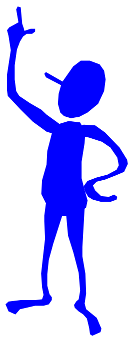
Impact of **pairwise comparisons**: $a \succsim b \Rightarrow a \succsim^N b$ and $a \succ b \Rightarrow$ *not* $b \succsim^P a$

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



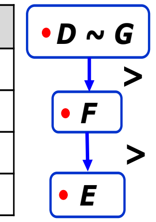
[29] Jeżeli dla danej pary zachodzi relacja konieczna, to zachodzi też relacja możliwa. Relacja konieczna jest zwrotna i przechodnia, a więc formalnie tworzy tzw. preporządek częściowy. Relacja możliwa jest zupełna, bo dla każdej pary zawsze zachodzi przynajmniej w jedną ze stron. Jest też negatywnie przechodnia, bo jeśli a nie jest możliwie preferowane nad b, zaś b nie jest możliwie preferowane nad c, to a nie jest też możliwie preferowane nad c. Jeżeli decydent powie, że a jest co najmniej tak dobre jak b, to znajdzie to odzwierciedlenie w relacji koniecznej. Jeśli stwierdzi, że a jest ściśle lepsze od b, to nie zajdzie relacja możliwa w drugą stronę. W przypadku braku jakiejkolwiek informacji preferencyjnej relacja konieczna sprowadza się do relacji dominacji, a możliwa jest relacją zupełną i zachodzi dla wszystkich par. Gdy informacja preferencyjna jest kompletna i ma postać rankingu wszystkich wariantów w zbiorze, to relacje konieczna i możliwa są sobie równe.

The Necessary Preference – First Iteration

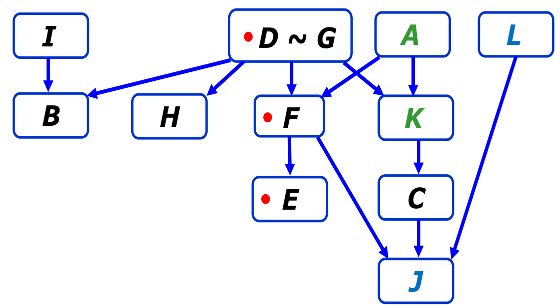
- **Necessary relation** can be presented using a Hasse diagram
- **Hasse diagram** is a type of mathematical diagram used to represent a finite partially ordered set in the form of a drawing of its **transitive reduction** ($A \succeq^N K$ and $K \succeq^N C \Rightarrow A \succeq^N C$)
- The necessary 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

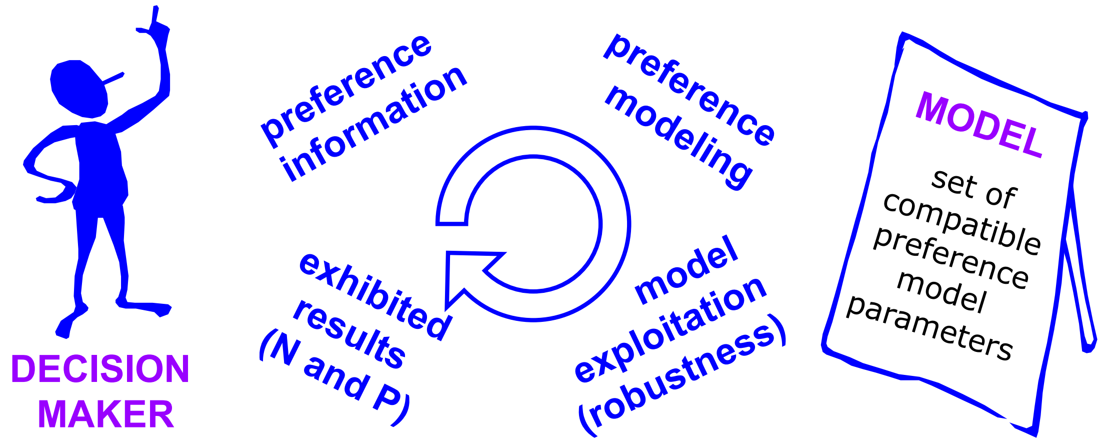
- Pairs such as (*I*,*B*), (*D*,*H*), and (*A*,*E*) are compared in the same way by all compatible value functions
- For pairs such as (*I*,*D*), (*H*,*F*), and (*L*,*K*), the results of a comparison is not univocal

[30] Dzięki przechodniości relacji koniecznej można ją przedstawić w postaci diagramu Hassego. Jest to graf, w którym pomijane są łuki wynikające z przechodniości relacji. Jeśli więc A jest koniecznie preferowane nad K, zaś K nad C, to A jest też koniecznie preferowane nad C, ale tego łuku nie rysujemy. Na slajdzie przedstawiłem diagram relacji koniecznej dla uprzednio dostarczonych preferencji. Zwróćcie uwagę, że odtwarza on mój ranking referencyjny, bo D jest koniecznie nierozróżnialne z G, są one koniecznie preferowane nad F, a F jest koniecznie preferowane nad E. Graf ten odtwarza też wszystkie relacje dominacji, jak choćby dla par (A,K) oraz (L,J). Poza tym zawiera też relacje, które wynikają z zastosowania wszystkich spójnych funkcji wartości. Analizując ten graf możemy wyróżnić pary, które są porównane w ten sam sposób przez wszystkie takie funkcje, jak choćby (I,B), (D,G) czy (A,E). Ale są też pary, takie jak (I,D) lub (H,F), dla których relacja konieczna nie zachodzi. Oznacza to, że dla niektórych funkcji lepszy jest jeden wariant, a dla innych lepszy jest drugi.

Learning Process

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

[31] Podejścia dezagregacji preferencji oferują mechanizm wzajemnego uczenia się metody i decydenta. Z jednej strony, metoda uczy się parametrów modelu preferencji, które są zgodne z polityką decyzyjną użytkownika. Z drugiej strony, decydent uczy się na podstawie konsekwencji zastosowania modelu na zbiorze wariantów. Analiza relacji koniecznej jest świetną podstawą dla potencjalnego wzbogacenia lub zmiany informacji preferencyjnej. W szczególności pary, które nie są w relacji koniecznej są idealnymi kandydatami do tego, żeby na ich temat wypowiedział się decydent w kolejnej iteracji.

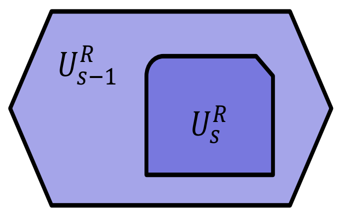
Incremental Set of Pairwise Comparisons

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)



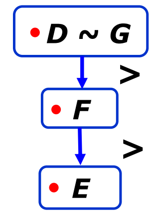
- **The necessary relation is enriched**: $\succsim_{s-1}^N \subseteq \succsim_s^N$
- What is certainly true becomes richer
- **The possible relation is impoverished**: $\succsim_{s-1}^P \supseteq \succsim_s^P$
- What is possibly true becomes narrower

[32] Rozważmy, więc takie zagnieżdżone zbiory porównań parami, powstające w ten sposób, że preferencje w poprzedniej iteracji są wzbogacone w iteracji kolejnej. Oczywiście nie ma sensu, by decydent wypowiadał się na temat par, dla których na pewnym etapie zachodzi już relacja konieczna. Jakie są konsekwencje wzbogacenia preferencji? Podstawową jest zawężenie zbioru spójnych funkcji wartości. A skoro ten zbiór jest coraz mniejszy i mniejszy, to relacja konieczna staje się coraz bogatsza, a relacja możliwa coraz bardziej uboga.

The Necessary Preference – Secod Iteration

1st iteration

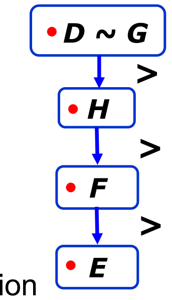
initial preference information



Reference alternatives

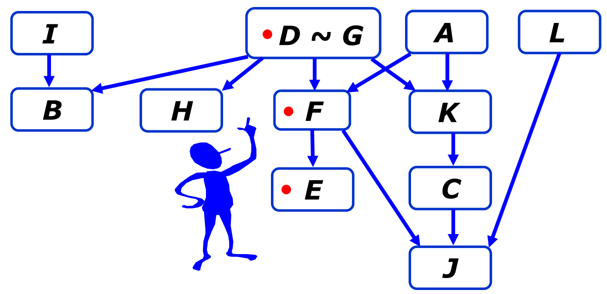
Alt.	$g_1\uparrow$	$g_2\uparrow$	$g_3\uparrow$
D	35	62	25
E	7	55	12
F	25	30	12
G	9	62	88
H	0	24	73

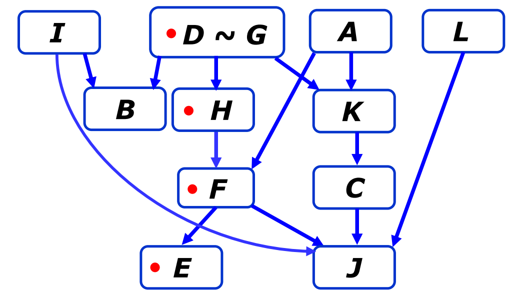
2nd iteration



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

[33] Rozważmy drugą iterację naszego problemu przez dodanie do rankingu referencyjnego wariantu H tak, by był on preferowany nad F. Powoduje to wzbogacenie relacji koniecznej. Część wynika z bezpośredniej konsekwencji dodanej informacji preferencyjnej, ale są też pary takie jak (I,J), dla których relacja zachodzi wskutek ograniczenia zbioru spójnych funkcji wartości. Taką interakcję można kontynuować do momentu, w którym uznamy, że relacja konieczna jest wystarczająco decyzyjna. Dodam tylko, że na Politechnice opracowaliśmy strategie aktywnego uczenia, które sugerują, które pary wariantów powinien porównywać decydent tak, by była z tego największa korzyść.

The UTA^{GMS} Method – Summary

Preference information

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

UTA^{GMS}

Preference model:

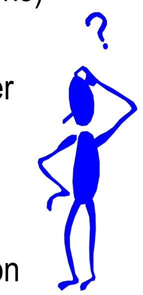
- A set of additive value functions with general marginal value functions (can be – and more often is – used with linear or piecewise linear functions)

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.

[34] Podsumowując, metoda UTA-GMS akceptuje pośrednią informację preferencyjną w postaci rankingu częściowego dla podzbioru wariantów referencyjnych, a więc porównań parami. Jako model preferencji wykorzystuje zbiór wszystkich spójnych addytywnych funkcji wartości z ogólnymi funkcjami cząstkowymi. Ten sam pomysł można jednak wykorzystać w kontekście funkcji liniowych lub odcinkami liniowych. Dalej programowanie liniowe jest wykorzystywane do oceny spójności dostarczonych preferencji oraz sprawdzenia prawdziwości dwóch relacji, koniecznej i możliwej, dla wszystkich par wariantów. Najbardziej wartościowym wynikiem jest relacja konieczna, która narzuca na zbiór wariantów porządek częściowy.

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:

$$U(\mathrm{x})-U(\mathrm{y}) \geq U(\mathrm{w})-U(\mathrm{z})$$

- ***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:

$$u_i(\mathrm{x})-u_i(\mathrm{y}) > u_i(\mathrm{w})-u_i(\mathrm{x})$$

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

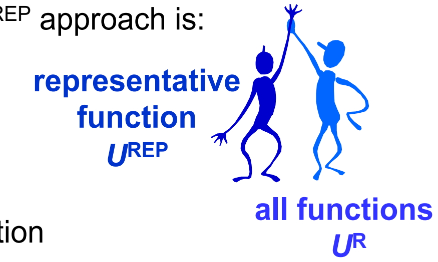
- „the increase of $g_i(a)=6$ by 3 is at least as attractive as the decrease of $g_j(a)=8$ by 2", which is represented by the LP constraint: $u_i(9)-u_i(6) \geq u_j(8)-u_j(6)$

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.*

[35] Metoda UTA-GMS została rozszerzona na wiele sposobów. Chciałbym teraz omówić dwa z takich rozszerzeń. Jednym z nich jest metoda GRIP. Posiada ona wszystkie cechy swojej poprzedniczki, ale dopuszcza też informację preferencyjną innego typu dotyczącą intensywności preferencji. Decydent może porównywać dwie pary wariantów holistycznie mówiąc, że wariant x jest preferowany nad wariant y bardziej, tak samo lub co najmniej tak bardzo jak wariant w nad wariant z. Te stwierdzenia są odpowiednio tłumaczone do ograniczeń liniowych. Mogą one też mieć charakter częściowy, odwołując się do jednego kryterium lub różnych kryteriów. Przykładowo, decydent może stwierdzić, że x jest preferowane nad y bardziej, tak samo lub co najmniej tak mocno jak w nad z, biorąc pod uwagę kryterium gj. Tłumacząc takie stwierdzenie, nie odwołujemy się do wartości globalnych, ale cząstkowych na danym kryterium.

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: $U(a) \geq U(b)+\varepsilon$

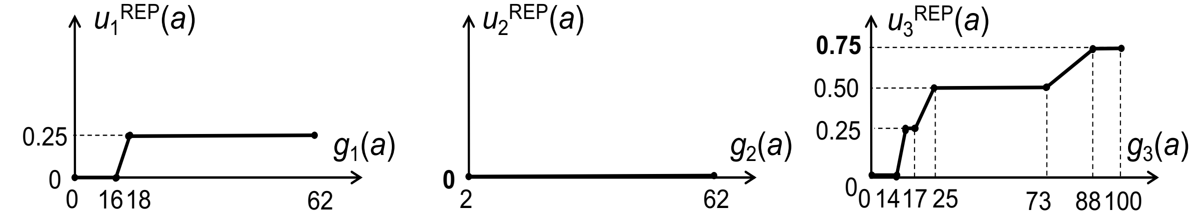
- To tie-breaking, one may **minimize** the difference of scores between alternatives (c,d) **not related** by the necessary preference ((*not* $c \succsim^N d$) and (*not* $d \succsim^N c$))

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

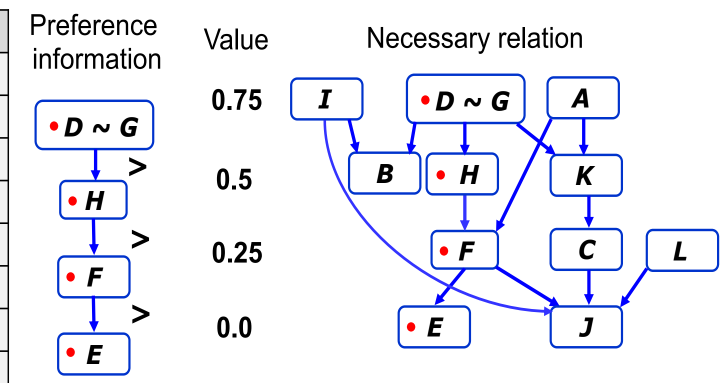
- **Maximize** $M\varepsilon - \delta$, where M is a "big value"

[36] Inne rozszerzenie dotyczy wyboru najbardziej reprezentatywnej funkcji; tu oznacza to funkcję najbardziej odporną. Chodzi o to, by mieć pojedynczy model, który będzie reprezentował wszystkie pozostałe, a te wszystkie pozostałe przyczynią się do jego wyboru. Pomysł jest następujący: dla par wariantów, dla których zachodzi relacji konieczna chcemy maksymalizować różnicę między ich wartościami globalnymi tak, by relacja preferencji stała się dla nich tak wyraźna jak tylko się da. Gdy to zrobimy, to dla par wariantów, dla których relacja konieczna nie zachodzi, chcemy minimalizować różnicę między ich wartościami globalnymi. Dlaczego? Ano skoro czasem lepszy jest jeden wariant, a czasem drugi, to w reprezentatywnym przypadku ich wartości powinny być bardzo zbliżone, jeśli nie takie same.

The Most Representative Value Function – Example



	$u_1(a)$	$u_2(a)$	$u_3(a)$	$U^{REP}(a)$	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

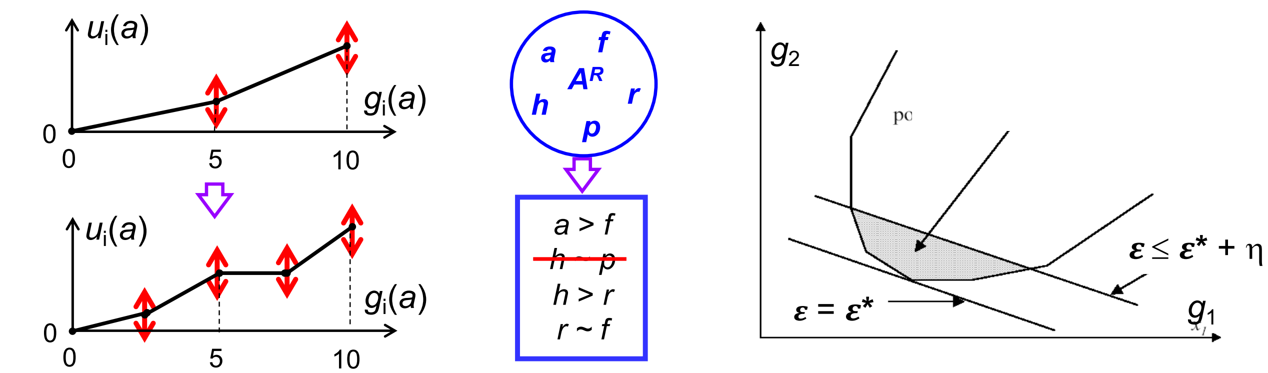
[37] Na slajdzie przedstawiłem funkcję wybraną w ten sposób dla preferencji z drugiej iteracji. Najbardziej różnicuje ona warianty na kryterium trzecim, a w ogóle ich nie różnicuje na kryterium drugim. Analiza takiej funkcji jest mniej abstrakcyjna niż całego zbioru nieskończenie wielu funkcji. Najważniejszy jest jednak reprezentatywny ranking wynikający z wartości globalnych. Solver przypisał je tak, by maksymalizować różnice między wariantami skojarzonymi relacją konieczną. Tu ta minimalna różnica jest równa 0.25. Poza tym wiele par wariantów nieporównywalnych względem koniecznej relacji uzyskało takie same wartości. W pewnym sensie taka reprezentatywna funkcja spłaszcza konsekwencje zastosowania wszystkich funkcji, pozwalając lepiej zrozumieć ranking konieczny.

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)



[38] Do tej pory zakładaliśmy, że informacja preferencyjna jest spójna z założonym modelem. Jednak nie zawsze musi tak być. Co w takim wypadku powinniśmy zrobić? Istnieją trzy możliwe ruchy. Pierwszy już znacie, bo polega on na zwiększeniu liczby liniowych odcinków dla funkcji wartości cząstkowych, o ile od razu nie były wykorzystywane funkcje ogólne. Dzięki temu model staje się bardziej elastyczny i może lepiej dopasować się do preferencji. Po drugie, możemy poszukać modelu, który jest suboptymalny pod względem pewnego błędu, choćby epsilona maksymalizowanego w regresji porządkowej. Jeśli nie da się zapewnić pełnej spójności, to chcemy by była ona tak wysoka jak się da. W końcu możemy też przyjąć, że popełniono jakiś błąd, albo stwierdzenia dotyczące preferencji są niespójne ze sobą, albo nie są zgodne z modelem addytywnym. Wtedy powinniśmy zmienić informację preferencyjną, eliminując lub zmieniając niektóre jej elementy.

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

$U(c) = U(d) \Rightarrow$

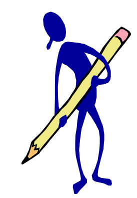
$U(c) \geq U(d)$

$U(d) \geq U(c)$

$\Downarrow$

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

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



[39] Skupimy się teraz na tym ostatnim sposobie. Pomysł polega na znalezieniu minimalnego podzbioru ograniczeń, które należy usunąć, aby przywrócić spójność. Ograniczenia definiujące model addytywny nie mogą być zmienione ani wyeliminowane. Dlatego skupiamy się na ograniczeniach związanych z porównaniami parami. Stają się one kandydatami do eliminacji. Z każdym porównaniem kojarzymy unikalną zmienną binarną. Tłumaczące je ograniczenia są przepisywane w specyficzny sposób. Musimy zagwarantować odtworzenie pożądanej relacji, gdy zmienna jest równa 0. Natomiast gdy zmienna jest równa 1, ograniczenie powinno być zawsze spełnione, będąc wyeliminowanym. Zauważcie, że w przypadku nierozróżnialności musimy najpierw przekształcić równość w dwie nieostre nierówności, a następnie przepisać je przy użyciu tej samej zmiennej binarnej.

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

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

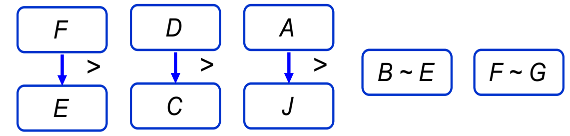
subject to

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

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

C (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 $B \sim E$



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

$$\begin{array}{l} U(F) > U(E) - v_{F,E} \\ U(D) > U(C) - v_{D,C} \\ U(A) > U(J) - v_{A,J} \\ U(B) \geq U(E) - v_{B,E} \\ U(E) \geq U(B) - v_{B,E} \\ U(F) \geq U(G) - v_{F,G} \\ U(G) \geq U(F) - v_{F,G} \end{array}$$

constraints related to 5 pairwise comparisons rewritten using the binary variables

$$v_{F,E},\ v_{D,C},\ v_{A,J},\ v_{B,E},\ v_{F,G} \in \{0,1\}$$

C (norm, monot, non-negat)

[40] Chcemy zachować jak najwięcej porównań. Jest to równoważne minimalizacji liczności podzbioru porównań parami, które muszą być usunięte, aby przywrócić spójność. To z kolei można osiągnąć minimalizując sumę zmiennych binarnych, z których każda jest związana z pojedynczym porównaniem. W przykładzie podanym po prawej stronie rozważamy pięć porównań parami, przy czym dwa z nich dotyczą nierozróżnialności. W ograniczeniach relacje między tymi parami wariantów są przepisywane z pomocą zmiennych binarnych, a wszystkie pozostałe ograniczenia dotyczące normalizacji, monotoniczności czy nieujemności pozostają nietknięte. Zmienne binarne, które w rozwiązaniu optymalnym przyjmują wartość 1, odpowiadają porównaniom parami, które trzeba usunąć. W przykładzie podanym na dole slajdu zakładam, że są dwie takie zmienne. Preferencja D nad C neguje relację dominacji, a nierozróżnialności są bardzo rygorystyczne i ciężko spełnić jednocześnie kilka takich stwierdzeń.

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*

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

- 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* ~ *E*

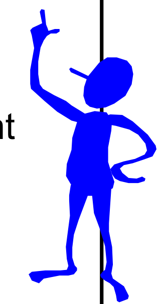
$$\boxed{v_{D,C} + v_{B,E} \leq 2 - 1}$$

- 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* ~ *G*

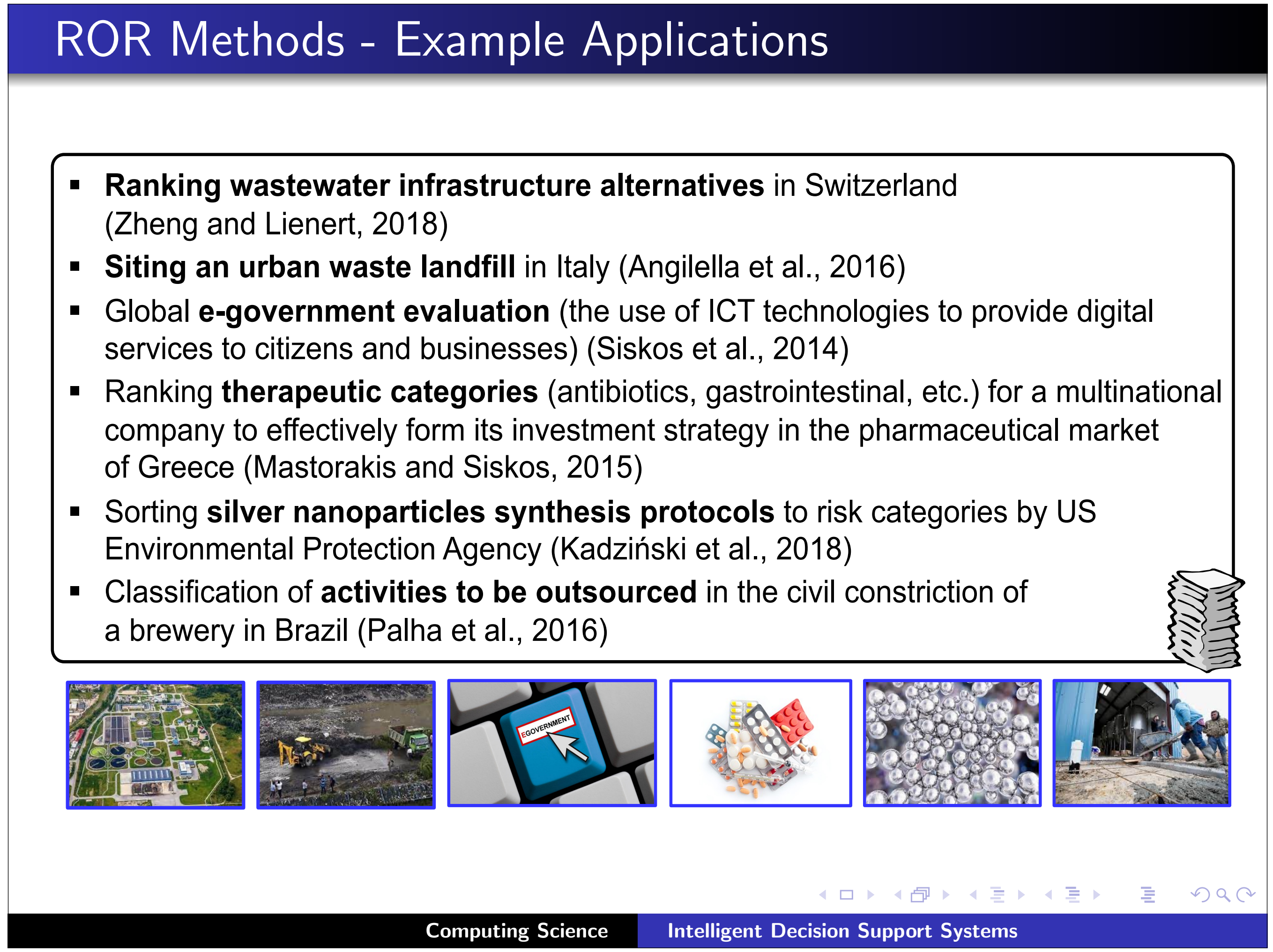
[41] Kompleksowe rozwiązanie niespójności polega na odkryciu wszystkich minimalnych podzbiorów leżących u jej podstaw. Powinny one zostać przedstawione decydentowi, aby mógł wybrać, które porównania parami usunąć. Aby znaleźć kolejne takie podzbiory, musimy postępować iteracyjnie. To znaczy, w każdej iteracji rozwiązujemy ten sam problem, ale wzbogacamy go o nierówności, które nie pozwalają na ponowne znalezienie tych samych rozwiązań, co w poprzednich iteracjach. Można to osiągnąć poprzez nałożenie ograniczenia na sumę zmiennych binarnych, które w poprzedniej iteracji przyjęły razem wartość 1. Suma taka musi być mniejsza lub równa liczbie tych zmiennych pomniejszonej o 1. W ten sposób nie mogą one wszystkie ponownie przyjąć wartości 1. Dla naszego przykładu oznaczałoby to, że suma zmiennych skojarzonych z parami (D,C) oraz (B,E) musi wynosić co najwyżej 1. Jeśli solver może znaleźć inne rozwiązanie, to przypisze jedynki zmiennym binarnym odpowiadającym innemu minimalnemu podzbiorowi ograniczeń. Jeśli nie, to problem będzie sprzeczny, a procedura może zostać zakończona.

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 21st 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



[42] Podejścia oparte na funkcji wartości należą do najpopularniejszych metod wspomagania decyzji, głównie ze względu na intuicyjną interpretację i łatwość obliczeń. Skupiliśmy się na regresji porządkowej, która uczy się z przykładów decyzji podobnie jak robią to metody w wielu innych poddziedzinach sztucznej inteligencji. Dlatego też metody te mają ścisłe powiązania ze statystycznym uczeniem maszynowym i odkrywaniem wiedzy. We wspomaganiu decyzji dezagregacji preferencji jest obecnie jednym z głównych nurtów i podobne podejścia opracowano dla różnych rodzin metod, w tym PROMETHEE czy ELECTRE. Jeśli chodzi o podejścia z rodziny UTA, to są one szeroko stosowane w praktyce. Główne dziedziny zastosowań to finanse, energia, marketing, zarządzanie środowiskiem oraz ocena projektów.



[43] Ostatni slajd dotyczy przykładowych, rzeczywistych zastosowań odpornej regresji porządkowej. W Szwajcarii korzystano z niej do wyboru infrastruktury oczyszczania ścieków, a we Włoszech do wyboru miejsca budowy wysypiska śmieci. Podejścia te stosowano też w tak nietypowych problemach jak ocena stopnia rozwoju usług cyfrowych w różnych państwach oraz tworzenie rankingu różnych rodzajów leków dla firmy farmaceutycznej, która chciała rozwijać się na rynku. Wreszcie istnieją też metody oparte na zasadzie odpornej regresji porządkowej, które adresują problem sortowania. W tym kontekście przykładowe zastosowania dotyczyły sortowania protokołów syntezy nanocząstek srebra do kategorii ryzyka oraz klasyfikacji zadań przy budowie browaru w Brazylii tak, by określić ich priorytet dla realizacji przez zewnętrzne podmioty. Dziękuję.