„Akademia Innowacyjnych Zastosowań Cyfrowych (AI Tech)”, projekt finansowany ze środków Programu Operacyjnego Polska Cyfrowa POPC.03.02.00-00-0001/20

INTELIGENTNE SYSTEMY WSPOMAGANIA DECYZJI

Lab III – Zasada Odpornej Regresji Porządkowej

Spis treści

1. Cel zajęć

Zajęcia dotyczą zasady odpornej regresji porządkowej (ang. *robust ordinal regression*). Omówimy dwa narzędzia elicytacji informacji preferencyjnej: agregację / dezagrację oraz analizę odporności. W pierwszej kolejności skupimy się na eksploatacji przestrzeni wag kryteriów spójnych ze zbiorem nierówności. Następnie podobny pomysł zastosujemy do eksploatacji zbioru funkcji użyteczności (wartości) spójnych z holistycznymi porównaniami parami. Jest to podstawa działania metody UTA^{GMS}, która stanowi rozszerzenie metody UTA. Przedstawimy też jej rozwinięcia w formie metody GRIP oraz procedury wyboru reprezentatywnej funkcji wartości. Omawiane pomysły są na tyle uniwersalne, że można je przenieść na grunt innych metod wspomagania decyzji.

2. Opis problemu

Rozważmy następujący problem decyzyjny:

Michał opracował stronę *fun4all.com*, by sprzedawać w Internecie urządzenia typu odtwarzacze mp3, konsole do gier, itd., przeznaczone głównie dla młodszych klientów. Chce zoptymalizować zakres produktów, które będzie oferował na stronie, tak by było one jak najbardziej atrakcyjne dla klientów. Michał nie może sobie jednak pozwolić na to, by wybór był bardzo duży.

Aby zidentyfikować, które odtwarzacze mp3 mógłby sprzedawać, Michał zapłacił drobną kwotę agencji marketingowej, która oceniła jakość odtwarzaczy z udziałem panelu potencjalnych młodych kupców. Każdy odtwarzacz był oceniony na trzech kryteriach (pojemność (g_1), autonomia (g_2), ergonomia/design (g_3)) na skali [0,100], gdzie 0 jest oceną najgorszą, a 100 - oceną najlepszą. Oceny sześciu odtwarzaczy oznaczonych symbolami a_1 - a_6 przedstawiono w poniższej tabeli.

	g_1	g_2	g_3
a_1	10	50	70
a_2	34	56	84
a_3	40	90	45
a_4	30	10	70
a_5	60	80	45
a_6	49	56	54

Aby ocenić ogólną atrakcyjność odtwarzaczy Michał, chce utworzyć ich ranking z wykorzystaniem prostego modelu preferencji, tj. sumy ważonej, gdzie wagi są nieujemne, a suma wag jest znormalizowana tak, by dawała 1:

$$U(a_i) = w_1 \cdot g_1(a_i) + w_2 \cdot g_2(a_i) + w_3 \cdot g_3(a_i), \text{ przy czym } w_1 + w_2 + w_3 = 1 \text{ oraz } w_i \geq 0,\ i = 1, 2, 3.$$

Aby zamodelować preferencje młodych klientów, Michał ma zamiar zadać swojemu bratankowi Antkowi kilka pytań. W odpowiedzi na te pytania, Antek przedstawił swoje preferencje w następujący sposób:

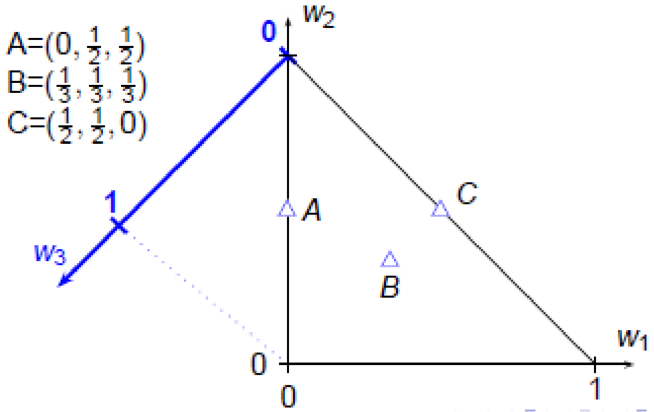
- „a_1 jest co najmniej tak atrakcyjny jak a_4”;
- „a_5 nie jest gorszy od a_3”;
- „a_2 jest co najmniej tak dobry jak a_6”.

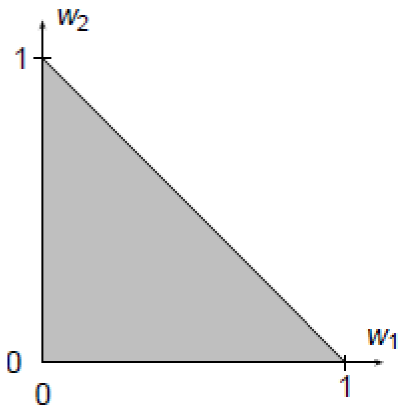
Dodatkowo, Michał uważa, że żadne kryterium nie powinno reprezentować więcej niż połowy wartości odtwarzacza (czyli nie powinno być ważniejsze niż dwa pozostałe kryterium razem wzięte).

Postarajmy się zamodelować preferencje wynikające z ograniczeń modelu (nieujemność wag i normalizacja) oraz preferencji przedstawionych przez Antka i Michała.

3. Ograniczenie na zbiór dopuszczalnych wag wraz z przyrostem informacji preferencyjnej

Przestrzeń możliwych trzech wag przed rozważaniem preferencji jest przestrzenią wszystkich wektorów wag [w_1, w_2, w_3] takich że $w_{1,2,3} \geq 0$ (nieujemność) oraz $w_1 + w_2 + w_3 = 1$ (normalizacja). Dla wygody reprezentacji takiej przestrzeni można ją narysować w dwóch wymiarach (w_1 oraz w_2), gdzie w_3 jest obliczana jako $1 - w_1 - w_2$, a w_1 oraz w_2 można odczytać jako współrzędne danego punktu (patrz dwa poniższe rysunki).



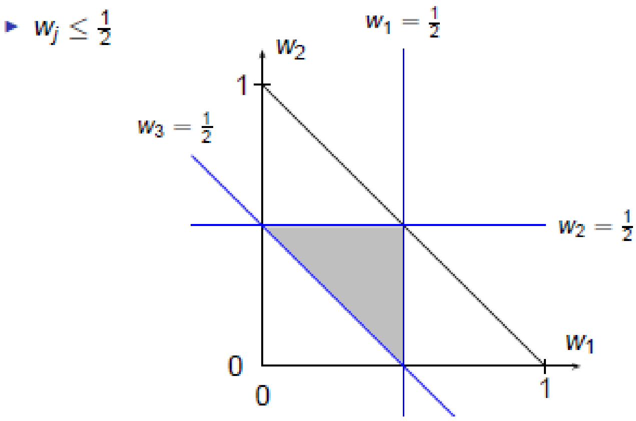


Michał uważa, że żadne kryterium nie powinno reprezentować więcej niż połowy wartości odtwarzacza, co przekłada się na ograniczenie:

waga każdego kryterium jest nie większa niż 1/2

$w_{1,2,3} \leq 1/2$

Przestrzeń możliwych wag ulega ograniczeniu (patrz szary obszar):



Antek uważa, że: „a_1 jest co najmniej tak atrakcyjny jak a_4", czyli a_1 S a_4 (a_1 przewyższa a_4) wtw. gdy $U(a_1) \geq U(a_4)$ (poniżej przedstawiono modelowanie preferencji dla tej pary):

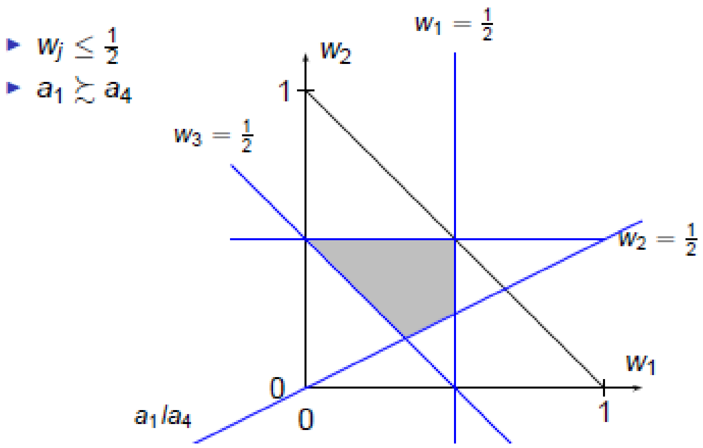
suma ważona dla a_1 = (10,50,70) ≥ suma ważona dla a_4 = (30,10,70)

$10w_1+50w_2+70w_3 \geq 30w_1+10w_2+70w_3$

$(50-10)w_2 \geq (30-10)w_1$

$2w_2 \geq w_1$

Przestrzeń możliwych wag ulega dalszemu ograniczeniu (patrz szary obszar):



Antek uważa, że: „a_5 nie jest gorszy od a_3”, czyli a_5 S a_3:

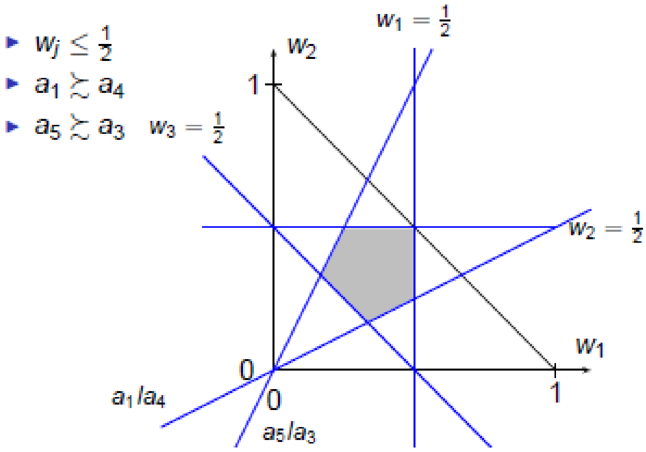
suma ważona dla a_5 = (60,80,45) ≥ suma ważona dla a_3 = (40,90,45)

$60w_1+80w_2+45w_3 \geq 40w_1+90w_2+45w_3$

$(60-40)w_1 \geq (90-80)w_2$

$2w_1 \geq w_2$

Przestrzeń możliwych wag ulega dalszemu ograniczeniu (patrz szary obszar):



Antek uważa, że: „a_2 jest co najmniej tak dobry jak a_6”, czyli a_2 S a_6

suma ważona dla a_2 = (34,56,84) ≥ suma ważona dla a_6 = (49,56,54)

$34w_1+56w_2+84w_3 \geq 49w_1+56w_2+54w_3$

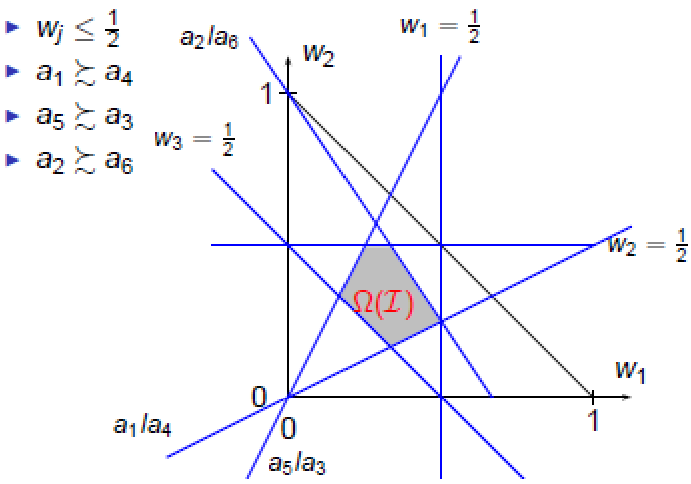
$(84-54)w_3 \geq (49-34)w_1$

$2w_3 \geq w_1$

$2(1-w_1-w_2) \geq w_1$

$2-2w_1-2w_2 \geq w_1$, a więc $2 \geq 3w_1+2w_2$

Przestrzeń możliwych wag ulega dalszemu ograniczeniu (patrz szary obszar):



Po zamodelowaniu wszystkich preferencji przestrzeń możliwych wag uległa znacznemu ograniczeniu w stosunku do przestrzeni pierwotnej. Pomimo tego, wektorów wag spójnych z preferencjami decydenta jest nieskończenie wiele.

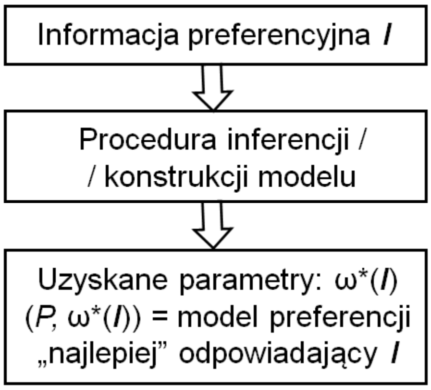
Pytanie: *jak w tym kontekście wypracować rekomendację dla decydenta?*

4. Narzędzia elicytacji informacji preferencyjnej

Podczas pierwszych zajęć wspomnieliśmy o trzech narzędziach elicytacji informacji preferencyjnej. Zostaną one szczegółowo omówione podczas dzisiejszych zajęć: dezagregacja-agregacja powiązana z wyborem reprezentatywnej instancji modelu preferencji spójnej z preferencjami decydenta, analiza odporności, a na końcu omówimy analizę niespójności.

Dezagregacja-agregacja i wybór jednego „reprezentatywnego" modelu

Pierwsze z narzędzi opiera się na arbitralnym wyborze reprezentatywnej spójnej instancji modelu preferencji. Narzędzie to znacie z metody UTA. Można je zrealizować na dwa sposoby: i) dać użytkownikowi możliwość wyboru takiej instancji z wykorzystaniem interaktywnego interfejsu użytkownika ograniczającego wybór tylko do instancji spójnych z preferencjami decydenta, ii) zastosować predefiniowaną procedurę wyboru, która w arbitralny sposób wybierze jedną spośród (nieskończenie) wielu instancji spójnych z preferencjami decydenta. W ramach dzisiejszych zajęć skupimy się na tym drugim sposobie i zaczniemy od najprostszej procedury, tj. wyboru najbardziej dyskryminującej instancji modelu, która "najlepiej" odzwierciedla preferencje.



Każde stwierdzenie dostarczone przez decydenta przekłada się na nierówność np.
$a_1 \; S \; a_4 \rightarrow U(a_1) - U(a_4) \geq 0$; $w_1 \leq 0.5$; itd. W ramach procedury wyznaczania najbardziej dyskryminującej instancji modelu te nierówności powinny być spełnione w najbardziej ewidentny sposób jak to tylko możliwe. W tym celu należy je przekształcić z wykorzystaniem nowej zmiennej (nazwijmy ją ε) tak, by jej maksymalizacja powodowało jak najsilniejsze spełnienie nierówności, np. $U(a_1) – U(a_4) \geq \varepsilon$ lub $\varepsilon + w_1 \leq 0.5$. Patrz poniższy program programowania liniowego:

Oryginalny zbiór ograniczeń na dozwolone (spójne) wartości parametrów modelu preferencji:

$U(a_1) – U(a_4) \geq 0$

$U(a_5) – U(a_3) \geq 0$

$U(a_2) – U(a_6) \geq 0$

$w_i \leq 0.5$, i=1,2,3

Problem programowania liniowego, który należy rozwiązać w celu wyboru reprezentatywnego (najbardziej dyskryminującego) modelu:

$\max \varepsilon$

p.o. $U(a_1) - U(a_4) \geq \varepsilon$

$U(a_5) - U(a_3) \geq \varepsilon$

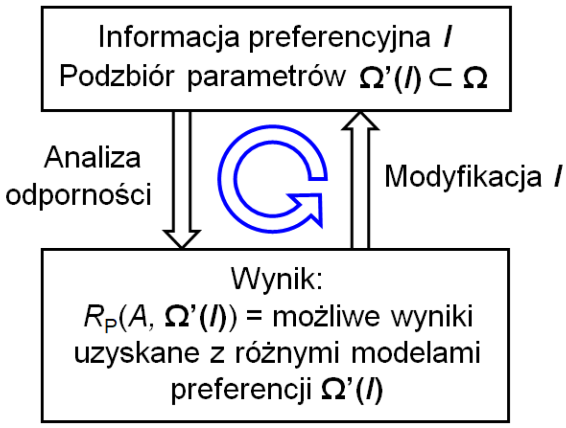
$U(a_2) - U(a_6) \geq \varepsilon$

$\varepsilon + w_i \leq 0.5$, $i=1,2,3$

$\omega \in \Omega(I)$ (normalizacja, nieujemność)

Analiza odporności

Analiza odporności polega na wzięciu pod uwagę wszystkich niepewności związanych z problemem decyzyjnym przy wypracowaniu rekomendacji. W naszym przypadku ta niepewność wiąże się z wyborem spójnej instancji modelu preferencji, a analiza odporności jest realizowana przez uwzględnienie wszystkich spójnych instancji modelu preferencji.



- **Główna trudność**: jak zdefiniować odporny ranking w kontekście wszystkich spójnych instancji modelu preferencji (w naszym przypadku - wszystkich spójnych wektorów wag)?
- Załóżmy, że *a* **jest „odpornie" = koniecznie preferowane** (w sensie słabym) nad *b* jeśli: $U(a) \geq U(b)$, $\forall\ \omega \in \Omega(I)$, tj. jeśli *a* jest co najmniej tak dobre jak *b* dla wszystkich wektorów wag spójnych z preferencjami decydenta;
- Aby zweryfikować, czy *a* jest odpornie preferowane nad *b*, musimy sprawdzić minimalną wartość różnicy $U(a) - U(b)$ przy ograniczeniach $\omega \in \Omega(I)$; oznaczmy ją przez $Min_{\omega \in \Omega(I)}(U(a) - U(b))$; jeśli $Min_{\omega \in \Omega(I)}(U(a) - U(b)) \geq 0$, to relacja konieczna dla pary (a,*b*) zachodzi;
- Alternatywnie można stwierdzić, że *a* jest **możliwie preferowane** (w sensie słabym) nad *b* jeśli: $U(a) \geq U(b)$, □$\omega \in \Omega(I)$, tj. jeśli *a* jest co najmniej tak dobre jak *b* dla co najmniej jednego wektora wag spójnego z preferencjami decydenta;

- Aby zweryfikować, czy *a* jest możliwie preferowane nad *b*, musimy sprawdzić maksymalną wartość różnicy U(*a*)−U(*b*) przy ograniczeniach $\omega \in \Omega(I)$; oznaczmy ją przez $Max_{\omega\in\Omega(I)}(U(a) - U(b))$; jeśli $Max_{\omega\in\Omega(I)}(U(a) - U(b)) \geq 0$, to relacja możliwa dla pary (a,*b*) zachodzi;
- Zgodnie z wartościami $Min_{\omega\in\Omega(I)}(U(a) - U(b)) \geq 0$ oraz $Max_{\omega\in\Omega(I)}(U(a) - U(b)) \geq 0$ można ustalić ranking konieczny i możliwy.

5. *UTA^GMS – zasada odpornej regresji porządkowej oraz porównanie z metodą UTA*

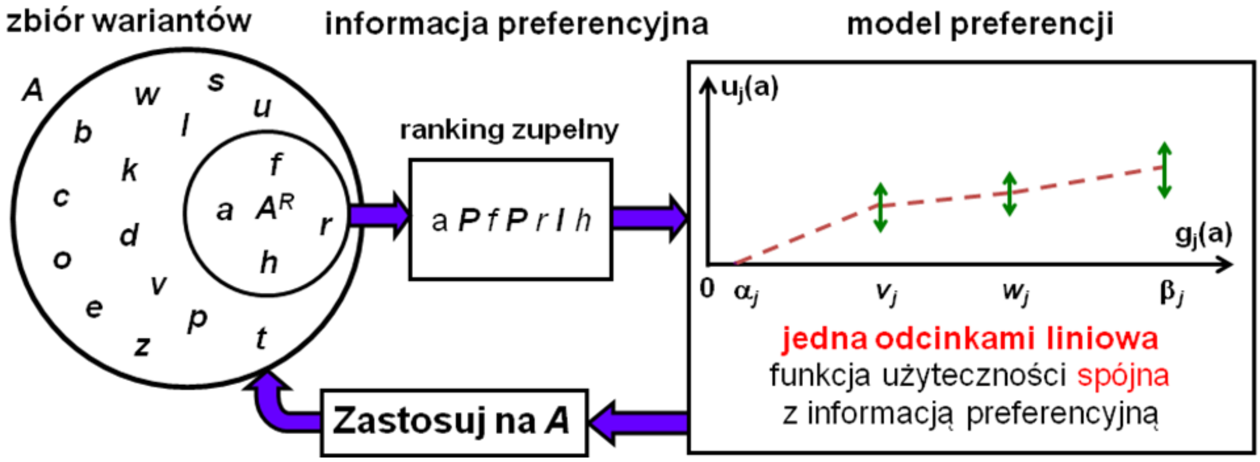
Cechy wspólne między UTA i UTA^{GMS}:

- model preferencji to addytywna funkcja użyteczności: $U(a) = \sum_{j=1,\dots,n} u_j(a)$;
- dezagregacja-agregacja (regresja porządkowa) – informacja preferencyjna dotycząca „fragmentu" końcowego wyniku jest podawana przez decydenta, a model lub modele nauczone przez metodę są stosowane, aby uzyskać ranking wszystkich wariantów.

Różnice pomiędzy metodami UTA oraz UTA^{GMS} omówiono ze względu na akceptowalną informację preferencyjną, wykorzystywany model preferencji oraz sposób wypracowania rankingu.

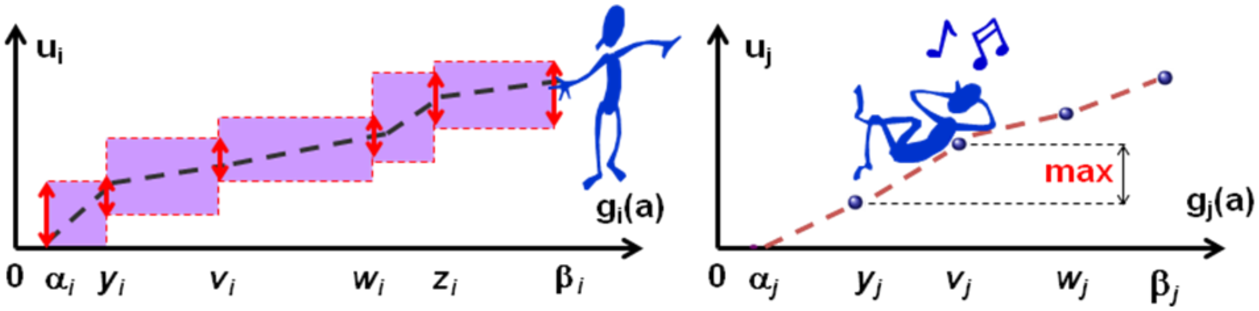
Metoda UTA

Informacja preferencyjna: porządek zupełny na zbiorze wariantów referencyjnych (warianty referencyjne muszą być ustawione w ranking, niedopuszczający nieporównywalności, np. a ***P*** f ***P*** r ***I*** h); dla każdego kryterium definiuje się też liczbę odcinków liniowych dla użyteczności cząstkowych (liczbę tzw. punktów charakterystycznych, w których funkcja użyteczności może zmienić swój kąt nachylenia względem osi x).



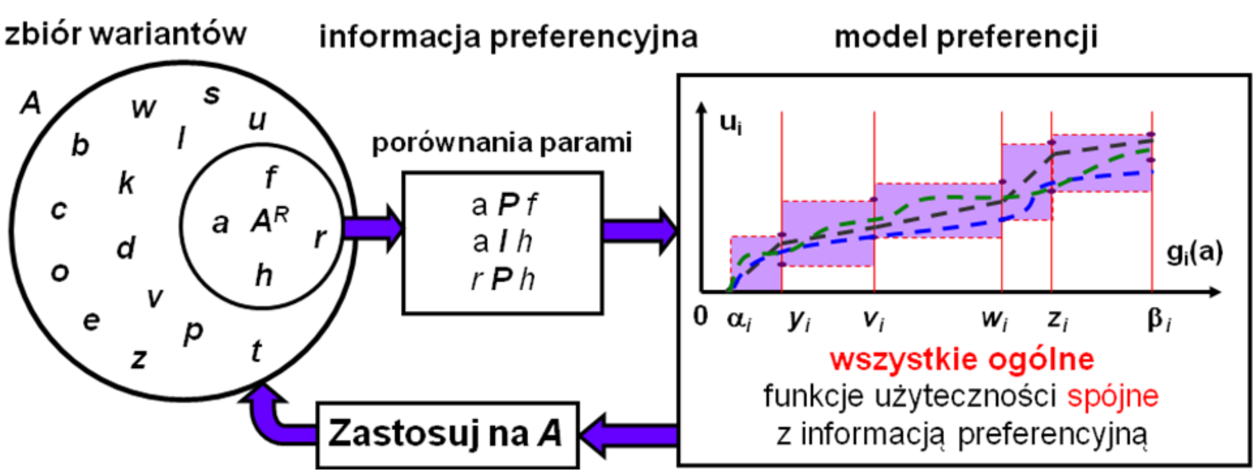
Model preferencji: funkcje cząstkowe odcinkami liniowe (ang. *piecewise linear*; należy zdefiniować liczbę punktów charakterystycznych; dla dwóch punktów charakterystycznych - funkcje są liniowe; zmiennymi w rozwiązywanym problemie są użyteczności w punktach charakterystycznych, a dla ocen które wypadają pomiędzy nimi, obliczenie użyteczności jest dokonywane z użyciem interpolacji liniowej).

Wypracowanie rankingu: z wykorzystaniem jednej funkcji użyteczności spójnej z preferencjami decydenta, wybranej w sposób interaktywny lub zgodnie z pewną domyślną regułą.

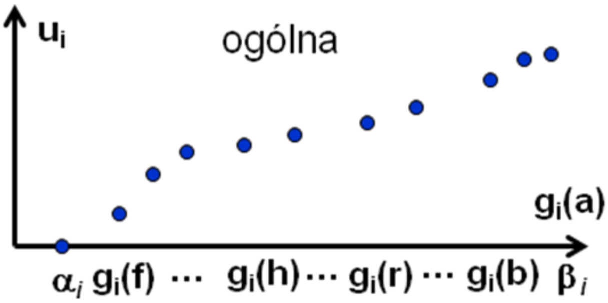


Metoda UTA^{GMS}

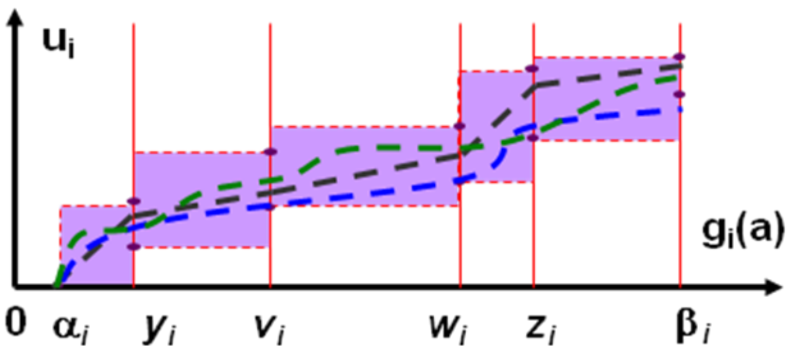
Informacja preferencyjna: porównania parami dla wariantów referencyjnych - porządek częściowy wariantów referencyjnych (można dostarczyć porównania parami, które de facto są rankingiem dopuszczającym nieporównywalności; np. a ***P*** f, a ***I*** h, r ***P*** h, ale choćby a oraz r nie są w żaden sposób porównane).



Model preferencji: funkcje cząstkowe monotoniczne ogólne (ang. *general*; punkty charakterystyczne odpowiadają wszystkim unikalnym ocenom wariantów; pomiędzy nimi przebieg jest nieistotny, ponieważ nie ma tam żadnych ocen); ***uwaga***: można podobny pomysł wykorzystać w kontekście funkcji liniowych lub odcinkami liniowych, ale oryginalnie w metodzie UTA^{GMS} zaproponowano wykorzystanie funkcji ogólnych, które nie wymagają podanie liczby segmentów lub punktów charakterystycznych.



Wypracowanie rankingu: wszystkie funkcje spójne z preferencjami decydenta; ranking w formie koniecznych i możliwych relacji preferencji, odzwierciedlających preferencję dla wszystkich lub co najmniej jednej funkcji użyteczności spójnej z preferencjami decydenta.



6. *Definicja zbioru addytywnych funkcji użyteczności (złożonych z monotonicznych funkcji użyteczności cząstkowych) spójnych z preferencjami decydenta*

Zbiór addytywnych funkcji użyteczności w metodzie UTA^{GMS} definiuje się poprzez zbiór ograniczeń liniowych wynikających z informacji preferencyjnej dostarczonej przez decydenta oraz założeń modelu (normalizacja globalnej/całkowitej użyteczności oraz monotoniczność cząstkowych funkcji użyteczności o charakterze ogólnym).

Informacja preferencyjna:

$U(a) \geq U(b) + \varepsilon$, gdy $a > b$ (a silnie preferowane nad b), gdzie $a, b \in A^R$

$U(a) = U(b)$, gdy $a \sim b$ (a nierozróżnialne z b), gdzie $a, b \in A^R$

normalizacja:

$u_j(x_j^1) = 0$, dla każdego kryterium $j \in F$, gdzie x_j^1 jest najgorszą oceną na danym kryterium

$\sum_{j \in F} u_j(x_j^{nj}) = 1$, gdzie x_j^{nj} jest najlepszą oceną na danym kryterium

monotoniczność:

$u_j(x_j^{k+1}) \geq u_j(x_j^k)$, dla każdego kryterium $j \in F$, oraz $k=1,\ldots,n_j-1$

gdzie $x_j^1 \le x_j^2 \le \ldots \le x_j^{nj-1} \le x_j^{nj}$ jest permutacją wszystkich unikalnych ocen wszystkich wariantów na kryterium $j \in F$ (przy czym n_j jest nie większe n – liczba wariantów; n_j jest mniejsze niż n, jeśli niektóre warianty mają taką samą ocenę). ε - arbitralnie mała wartość.

Uwaga: odniesienie w warunkach monotoniczności do użyteczności cząstkowych wszystkich unikalnych ocen wariantów pozwala uniknąć wielokrotnego przeformułowania problemów na późniejszych etapach działania metody; mimo to warto jednak wiedzieć, że w warunkach monotoniczności można odnieść się tylko do użyteczności cząstkowych dla wariantów referencyjnych (tj. tylko tych porównanych przez decydenta), bo tylko dla nich istnieją jakiekolwiek ograniczenia wynikające z podanej informacji preferencyjnej. Pozwala to na ograniczenie liczby zmiennych w problemie, ale wymaga unikalnego sformułowania przez uwzględnienie dodatkowych zmiennych przy sprawdzeniu prawdziwości relacji koniecznej i możliwej dla par wariantów, z których co najmniej jeden jest niereferencyjny.

Równoważne sformułowanie (różnica występuje tylko w warunkach monotoniczności):
Informacja preferencyjna:

$U(a) \ge U(b) + \varepsilon$, gdy $a > b$, gdzie $a, b \in A^R$

$U(a) = U(b)$, gdy $a \sim b$, gdzie $a, b \in A^R$

normalizacja:

$u_j(\alpha_j) = 0$, dla każdego kryterium $j \in F$, gdzie α_j jest najgorszą oceną na danym kryterium

$\sum_{j \in F} u_j(\beta_j) = 1$, gdzie β_j jest najlepszą oceną na danym kryterium

monotoniczność:

$u_j(x_j^{k+1}) \ge u_j(x_j^k)$, dla każdego kryterium $j \in F$, oraz $k=1,\ldots,n_j^R-1$

gdzie $x_j^1 \le x_j^2 \le \ldots \le x_j^{njR-1} \le x_j^{njR}$ jest permutacją wszystkich unikalnych ocen tylko wariantów referencyjnych na kryterium $j \in F$ (przy czym n_jR jest na pewno nie większe niż $|A^R|$ – liczba wariantów referencyjnych; może być mniejsze jeśli niektóre warianty referencyjne mają taką samą ocenę).

$u_j(x_j^1) \ge u_j(\alpha_j)$, dla każdego kryterium $j \in F$

$u_j(x_j^{njR-1}) \le u_j(\beta_j)$, dla każdego kryterium $j \in F$

7. *Relacje koniecznej i możliwej preferencji*

Oznaczmy przez $U(A^R)$ zbiór wszystkich funkcji użyteczności spójnych z preferencjami decydenta w postaci porównań parami wariantów referencyjnych ze zbioru A^R.

- Wariant a jest **koniecznie preferowany** (w sensie słabym) nad wariant b (oznaczenie **a $\ge^N$ b**), jeśli $U(a) \ge U(b)$ (użyteczność wariantu *a* jest nie mniejsza niż użyteczność wariantu *b*) dla **wszystkich** funkcji użyteczności w zbiorze $U(A^R)$ spójnych z preferencjami decydenta.

Oznaczmy przez d(a,b) = $Min_{U \in U(AR)} U(a) - U(b)$ minimalną różnicę użyteczności dla pary wariantów a oraz b. Relacja konieczna a $\geq^N$ b jest prawdziwa, gdy d(a,b) ≥ 0.

- Wariant a jest **możliwie preferowany** (w sensie słabym) nad wariant b (oznaczenie **a $\geq^P$ b**), jeśli $U(a) \geq U(b)$ dla **co najmniej jednej** funkcji użyteczności w zbiorze $U(A^R)$ spójnej z preferencjami decydenta.

  Oznaczmy przez D(a,b) = $Max_{U \in U(AR)} U(a) - U(b)$ maksymalną różnicę użyteczności dla pary wariantów a oraz b. Relacja możliwa a $\geq^P$ b jest prawdziwa, gdy D(a,b) ≥ 0.

Własności koniecznych oraz możliwych relacji preferencji:

- Relacja konieczna $\geq^N$ to pre-porządek częściowy (ang. *partial pre-order*) (relacja zwrotna i przechodnia);
- Relacja możliwa $\geq^P$ jest silnie zupełna (zawsze zachodzi a $\geq^P$ b lub b $\geq^P$ a), ale w ogólności nie jest przechodnia (tzn. nie zawsze a $\geq^P$ b oraz b $\geq^P$ c determinuje a $\geq^P$ c);
- Przy braku informacji preferencyjnej: relacja konieczna = relacja dominacji;
- Przy braku informacji preferencyjnej: relacja możliwa jest zupełna = wszystko jest możliwe;
- Przy wszystkich parach wariantów porównanych przez decydenta:
  ranking zupełny = ranking możliwy = ranking konieczny;
- Jeśli a $\geq^N$ b to a $\geq^P$ b (prawdziwość relacji koniecznej determinuje prawdziwość relacji możliwej);
- W ogólności nie jest możliwe odtworzenie jednego rankingu na podstawie drugiego (relacje konieczne i możliwe nie są dualne), tzn. a $\geq^N$ b nie wyklucza b $\geq^P$ a, bo U(a) może być ≥ U(b) dla wszystkich spójnych funkcji użyteczności U, a pomimo to U(b) może być ≥ U(a) dla co najmniej jednego U.

8. *Metoda GRIP*

GRIP rozszerza UTA^{GMS} poprzez wprowadzenie dodatkowego typu informacji preferencyjnej, która dotyczy tzw. **intensywności preferencji**:

- całościowe porównania intensywności preferencji dla par wariantów:
  np. „a jest (słabo) preferowane nad b co najmniej tak mocno jak c nad d”,
  co oznacza się przez $(a,b) \geq^* (c,d)$;
- cząstkowe porównania intensywności preferencji dla par wariantów na konkretnym kryterium:
  np. „a jest słabo preferowane nad b co najmniej tak mocno jak c nad d na kryterium g_j”,
  co oznacza się przez $(a,b) \geq_j^* (c,d)$.

Zamiast słabej preferencji, można odnieść się do silnej preferencji (oznaczanej przez >) lub relacji nierozróżnialności (oznaczanej przez ~).

Dodatkowo w metodzie GRIP sprawdza się też prawdziwość koniecznych i możliwych relacji intensywności preferencji (choć jest to rzadko wykorzystywane w praktyce), np.:

- $(a,b) \geq^{*,N} (c,d)$ jeśli U(a) – U(b) ≥ U(c) – U(d) dla wszystkich spójnych funkcji użyteczności,

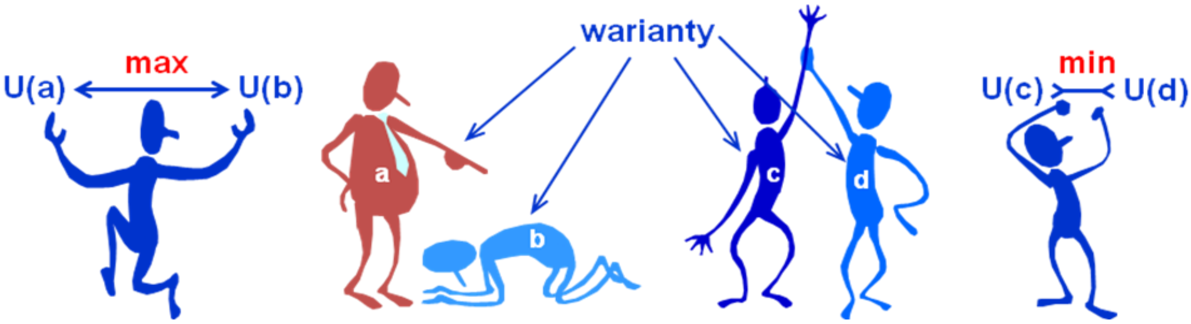
- $(a,b) \geq_j^{*,P} (c,d)$ jeśli $u_j(a) – u_j(b) \geq u_j(c) – u_j(d)$ dla co najmniej jednej spójnej funkcji użyteczności.

9. *Reprezentatywna funkcja użyteczności*

W niektórych sytuacjach analiza odporności może nie dostarczać wyników, które są wystarczająco decyzyjne. Przykładowo relacja konieczna może być zbyt rzadka, a relacja możliwa zbyt bogata, by realnie wspomóc podjęcie decyzji. Pojedyncza reprezentatywna funkcja użyteczności dostarcza takich precyzyjnych wyników. Nie chcielibyśmy jednak tracić wyników analizy odporności, więc powstał pomysł, aby to na nich oprzeć bardziej złożone kryteria, które posłużą do wyboru jednej reprezentatywnej funkcji użyteczności.

Zasada wyboru reprezentatywnej funkcji użyteczności w duchu odpornej regresji porządkowej:

- "Jeden za wszystkich, wszyscy za jednego”, tj. jedna funkcja reprezentuje wszystkie pozostałe, a wszystkie pozostałe mają wkład w definicję i wybór tej reprezentatywnej;
- Pomysł polega na podkreśleniu przewagi jednych wariantów nad innymi potwierdzonej przez wszystkie spójne funkcje użyteczności (prawdziwość relacji koniecznej potwierdzona przez wszystkie spójne funkcje użyteczności) oraz podkreślenie niejednoznaczności we wskazaniu lepszego wariantu dla innych par wariantów (brak prawdziwości relacji koniecznej w obie strony, tj. nieporównywalność względem relacji koniecznej oznaczająca, że dla co najmniej jednej funkcji użyteczności jeden wariant jest silnie preferowany nad drugi, a innej funkcji użyteczności - kolejność jest odwrotna).



Punkt wyjścia: zbiór ograniczeń definiujących funkcje użyteczności spójne z preferencjami decydenta oraz przeprowadzono analizę koniecznych $\geq^N$ oraz możliwych $\geq^P$ relacji preferencji.

10. **Procedura wyboru reprezentatywnej funkcji użyteczności** - chcemy zoptymalizować dwa cele w porządku leksykograficznym:

Do zbioru ograniczeń definiujących funkcje użyteczności spójne z preferencjami decydenta, dla par wariantów (a,b) takich, że a $\geq^N$ b , ale nie b $\geq^N$ a, dodaj następujący warunek do zbioru ograniczeń:

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

Maksymalizuj ε , co pozwala na podkreślenie różnicy użyteczności dla par wariantów, dla których zachodzi konieczna preferencja.

Dodaj warunek $\varepsilon=\varepsilon^*$, gdzie $\varepsilon^*=\max \varepsilon$ *do zbioru ograniczeń* (utrzymanie minimalnej różnicy na zoptymalizowanym poziomie).

Dla par wariantów (c,d) takich, *że* ani c $\geq^N$ d ani d $\geq^N$ c, dodaj następujące warunki do zbioru ograniczeń:

$$U(c)-U(\mathrm{d}) \leq \delta \quad \text{oraz} \quad U(d)-U(c) \leq \delta,$$

co odpowiada $|U(c)-U(\mathrm{d})| \leq \delta$.

Minimalizuj δ, co pozwala na zmniejszenie różnica dla par wariantów, dla których nie zachodzi konieczna preferencja.

Reprezentatywna funkcja ma więc podkreślić ewidentną przewagę jednych wariantów nad innymi, która jest potwierdzona przez relację konieczną, oraz zniwelować taką różnicę dla par, dla których porównanie względem relacji koniecznej nie dało pozytywnych wyników (oznacza to, że dla co najmniej jednej funkcji jeden wariant ściśle lepszy niż drugi, a dla innej funkcji kolejność w rankingu jest odwrotna - wobec tej niejednoznaczności w wynikach analizy odporności, chcielibyśmy zminimalizować różnicę między użytecznościami takich wariantów w reprezentatywnym przypadku).

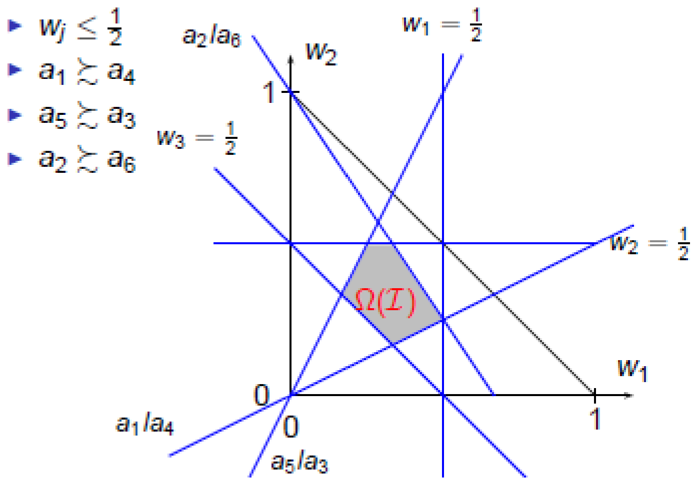
11. **Wykrywanie niespójności w informacji preferencyjnej**

- Kontynuacja procesu elicytacji informacji preferencyjnej dla problemu fun4all.com w kontekście wykrywania niespójności w informacji preferencyjnej;
- Niespójność pojawia się, gdy nie ma żadnej instancji modelu preferencji, która odtwarzała by wszystkie preferencje dostarczone przez decydenta. Z matematycznego punktu widzenia,

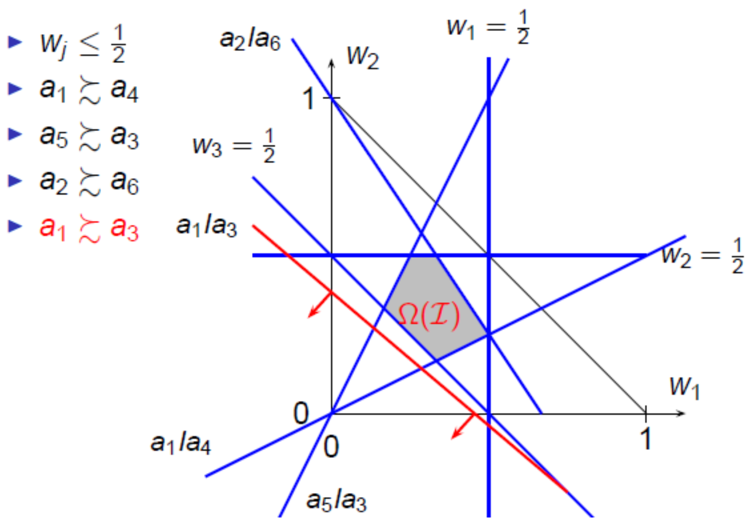
układ ograniczeń modelujący preferencje decydenta jest wtedy sprzeczny, a przestrzeń parametrów modelu preferencji spójnych z preferencjami decydenta pusta.

- Punktem wyjścia analizy jest sytuacja, w której wykorzystywanym modelem preferencji jest suma ważona oraz dostarczono następujących preferencji w postaci trzech porównań parami oraz bezpośrednich ograniczeń na wagi poszczególnych kryteriów:
  - „a_1 jest co najmniej tak atrakcyjny jak a_4”, „a_5 nie jest gorszy od a_3” oraz „a_2 jest co najmniej tak dobry jak a_6”;
  - „waga każdego kryterium nie jest większa niż 0.5".

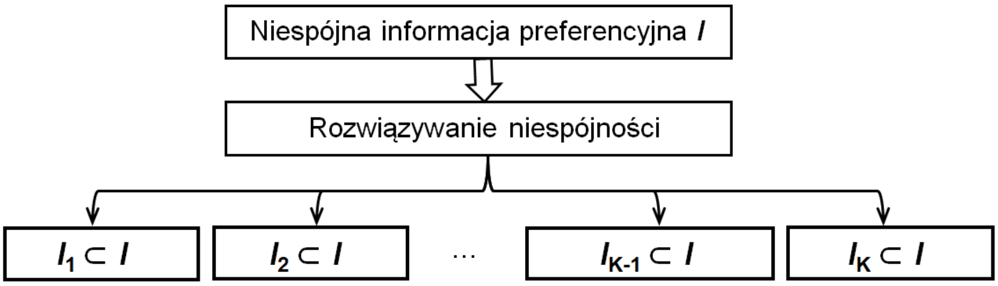
  Modelowanie preferencji doprowadziło do wyróżnienia niepustej przestrzeń wag odtwarzających preferencje użytkownika (patrz szary obszar):



- Załóżmy, że oprócz ograniczeń na wagi i trzech porównań parami doszło kolejne stwierdzenie, że "a_1 jest co najmniej tak dobre jak a_3" (patrz czerwona linia na poniższym rysunku). Doprowadza to do sytuacji, w której przestrzeń wag spójnych z preferencjami decydenta staje się pusta, co oznacza niespójność preferencji z założonym modelem preferencji.

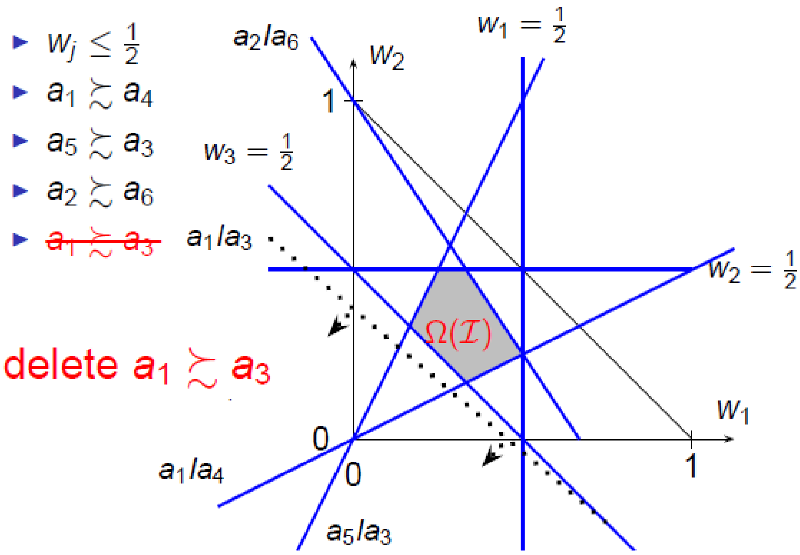


- **Wykrycie niespójności w informacji preferencyjnej i jej rozwiązanie**: należy pomóc decydentowi w przywróceniu spójności. Aby rozwiązać problem niespójności, należy zidentyfikować maksymalne podzbiory fragmentów informacji preferencyjnej, dla których zbiór instancji spójnych z preferencjami decydenta jest niepusty. Jest to problem równoważny znalezieniu minimalnych podzbiorów elementów informacji preferencyjnej I1, ..., IK, które należy usunąć lub zmienić, aby informacja stała się spójna. Takich podzbiorów może być wiele i powinno przedstawić się je decydentowi, aby mógł on wskazać który zostanie usunięty.

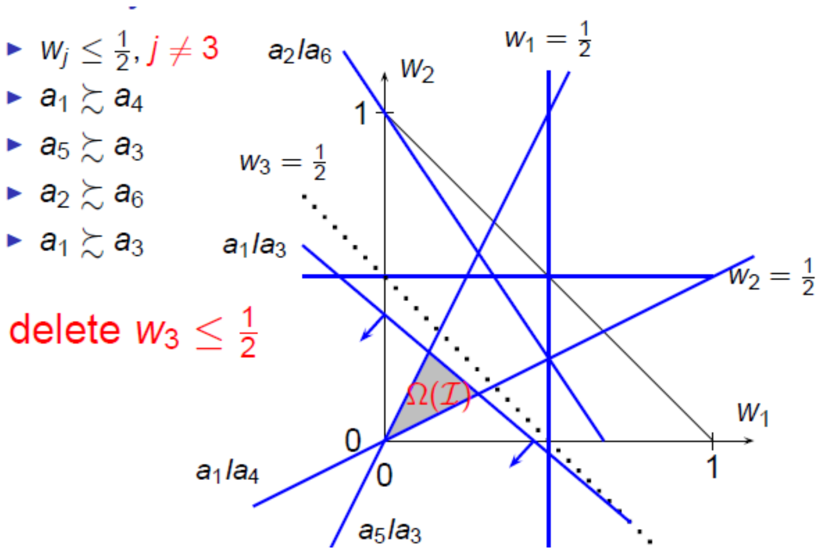


- Dla rozważanego problemu istnieją dwa minimalne sposoby przywrócenia spójności. Pierwszy z nich polega na usunięciu porównania "a_1 jest co najmniej tak dobre jak a_3" (a_1 *S* a_3), a drugi na usunięciu ograniczenia wartości wagi w_3. Przedstawiono je na poniższych rysunkach.

  Pierwszy minimalny podzbiór elementów informacji preferencyjnej, który należy usunąć, by przywrócić jej spójność z założonym modelem preferencji:



Drugi minimalny podzbiór elementów informacji preferencyjnej, który należy usunąć, by przywrócić jej spójność z założonym modelem preferencji:



Oryginalny zbiór ograniczeń na dozwolone wartości parametrów modelu preferencji:

$U(a_1) – U(a_4) \ge 0$

$U(a_5) – U(a_3) \ge 0$

$U(a_2) – U(a_6) \ge 0$

$U(a_1) – U(a_3) \ge 0$

$w_i \le 0.5$, i=1,2,3

Wprowadźmy binarne zmienne decyzyjne v_{ij} ($v_{ij} \in \{0,1\}$, dla i,j = 1,...,m) przyjmujące wartości 1 lub 0. Wartość $v_{ij} = 0$ oznacza zachowanie ograniczenia w niezmienionej formie, zaś 1 oznacza że równanie jest zawsze spełnione. W celu znalezienia minimalnego podzbioru niespójnych porównań parami należy przekształcić powyższe równania z zastosowaniem zmiennych binarnych.

Problem mieszanego całkowitoliczbowego programowania liniowego, (ang. *Mixed-Integer Linear Programming*) który należy rozwiązań w celu wyboru reprezentatywnego (najbardziej dyskryminującego) modelu:

$Min_{\omega \in \Omega(I)}(v_{14} + v_{53} + v_{26} + v_{14})$

$U(a_1) – U(a_4) + v_{14} \ge 0$

$U(a_5) – U(a_3) + v_{53} \ge 0$

$U(a_2) – U(a_6) + v_{26} \ge 0$

$U(a_1) – U(a_3) + v_{13} \ge 0$

$w_i \le 0.5$, i=1,2,3

$v_{ij} \in \{0,1\}$, dla i,j = 1,...,m.

Zauważmy iż ten model jest w stanie znaleźć jeden zbiór równań spójnych naraz na przykład v_{13}=1 oznaczające usunięcie porównania (a_1 *S* a_3). W celu znalezienia kolejnych podzbiorów spójnych należy zagwarantować iż suma zmiennych binarnych dla wcześniej znalezionych podzbiorów V_k jest równa k-1 gdzie k jest liczbą równań w V_k np:

$Min_{\omega \in \Omega(I)}(v_{14} + v_{53} + v_{26} + v_{14})$

$U(a_1) - U(a_4) + v_{14} \geq 0$

$U(a_5) - U(a_3) + v_{53} \geq 0$

$U(a_2) - U(a_6) + v_{26} \geq 0$

$U(a_1) - U(a_3) + v_{13} \geq 0$

$v_{13} = 0$

$w_i \leq 0.5$, i=1,2,3

$v_{ij} \in \{0,1\}$, dla i,j = 1,...,m.

Proces oddawania kolejnych równań wykonujemy aż do znalezienia sprzecznego układu równań który świadczy o znalezieniu wszystkich podzbiorów spójnych informacji preferencyjnych.

Fundusze Europejskie Polska Cyfrowa

Rzeczpospolita Polska

Unia Europejska Europejski Fundusz Rozwoju Regionalnego

„Akademia Innowacyjnych Zastosowań Cyfrowych (AI Tech)”, projekt finansowany ze środków Programu Operacyjnego Polska Cyfrowa POPC.03.02.00-00-0001/20