

„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 I – Wprowadzenie do Inteligentnych Systemów Wspomagania Decyzji

1. **Cel zajęć**

Zajęcia będą podzielone na dwie części. W pierwszej omówione zostaną wybrane pojęcia związane z wielokryterialnym wspomaganie decyzji. Pozwoli nam to na znalezienie wspólnego, bardziej formalnego języka, który będzie wykorzystywany podczas tego przedmiotu. Przedstawimy więc cel procesu decyzyjnego, głównych aktorów w niego zaangażowanych oraz podstawowe elementy rozważane i wykorzystywane w takich procesach. Główny nacisk zostanie położony na informację preferencyjną, związany z nią proces elicytacji (gromadzenia, pobierania), naturę tego procesu oraz wykorzystywane w nim narzędzia. Narzędzia te będą przedmiotem naszej szczegółowej analizy w ciągu kilku następnych tygodni.

Druga, zdecydowanie dłuższa część zajęć będzie dotyczyła przypomnienia podstawowych metod poznanych na przedmiocie Wspomaganie Decyzji, które będą stanowiły dla nas bazę do poznania bardziej zaawansowanych metod w dalszej części semestru. Dla niektórych z Was będzie to zupełnie nowy materiał. Celem jest więc wyrównanie Waszego poziomu. Powtórka ta ma charakter ćwiczeniowy (z wykorzystaniem dedykowanego arkusza ćwiczeń), a rozwiązywanie zadań jest w niej połączone z elementami teorii oraz szczegółowym omówieniem celu ćwiczeń, które można znaleźć w ostatnim rozdziale tego pliku prowadzącego.

2. **Podstawowe pojęcia związane z wielokryterialnym wspomaganie decyzji**

Proces decyzyjny:

- **Start:** określenie celu (lub celi) do osiągnięcia, które wymagają podjęcia decyzji; odpowiedź na pytania: jaki problem rozwiązujemy? jaki typ wyniku chcielibyśmy uzyskać?
- Identyfikacja czynników, mających wpływ na podjęcie decyzji; wśród nich można wyróżnić czynniki ograniczające naszą decyzję (np. dostępność pieniędzy czy lokalizacja), jak i te które mają wpływ na nasze postrzeganie alternatywnych sposobów rozwiązania problemu;
- Określenie zbioru decyzji dopuszczalnych; zwykle istnieje wiele (tj. więcej niż jeden) alternatywnych sposobów osiągnięcia celu;
- Zbudowanie modelu i określenie zbioru decyzji optymalnych (jeśli istnieją) z punktu widzenia naszego problemu;

- **Wspomaganie decyzji** ma na celu pomoc w uzyskaniu (elementów) odpowiedzi na pytania stawiane przez osoby zaangażowane w proces decyzyjny, a także w uzasadnieniu decyzji oraz rekomendacji zachowania, które jest najbardziej spójne z celami i systemem wartości uczestnika/ów procesu;
- **Zakończenie:** podjęcie ostatecznej decyzji przy potencjalnym wzięciu pod uwagę rekomendacji wypracowanej w procesie wspomagania decyzji.

Podstawowi aktorzy zaangażowani w proces decyzyjny:

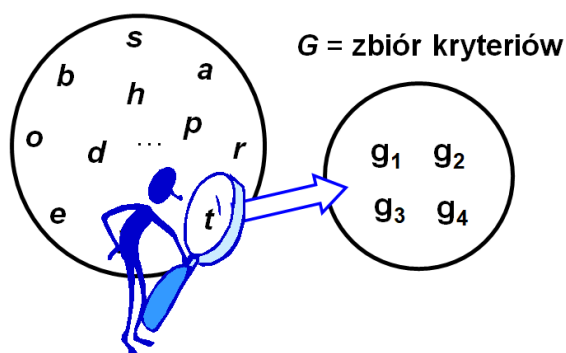
☀ **Decydent:** podmiot (człowiek, grupa, maszyna), w imieniu którego lub dla którego realizowane jest wspomaganie decyzji; to on ma podjąć ostateczną decyzję.

☀ **Analitik:** pełni funkcję pomocniczą wobec decydenta; ułatwia postęp procesu wspomagania decyzji poprzez interakcję z decydentem. Zadaniem tego przedmiotu jest wykształcenie w studentach podstawowych umiejętności, które powinien posiadać analitik.

Przedmiot rozważań w ramach procesu decyzyjnego:

☀ **Zbiór wariantów:** zbiór czynności, obiektów decyzji, które potencjalnie można zrealizować i wobec których osoby zaangażowane w proces przejawiają zainteresowanie. Zbiór wariantów najczęściej oznacza się dużą literą A: $A = \{a_1, a_2, \dots, a_m\}$ – np. pacjenci, wnioski kredytowe, kandydaci na określone stanowisko, projekty, itd.

$A = \text{zbiór wariantów}$



W procesie decyzyjnym czynniki niezależne od decydenta zazwyczaj są **warunkami ograniczającymi decyzję**, a zależne – **kryteriami** oceny.

☀ Kryterium:

Funkcja (najczęściej rzeczywista) g_j zdefiniowana na zbiorze wariantów A , odzwierciedlająca wartość wariantu z określonego punktu widzenia w ten sposób, że aby porównać parę wariantów $a, b \in A$ pod tym względem, wystarczy porównać wartości $g_j(a)$ oraz $g_j(b)$. Rodzinę (zbiór) kryteriów najczęściej oznacza się dużą literą "G" lub "F": $G = \{1, 2, \dots, n_{\text{kryt}}\}$.

Rodzina kryteriów musi spełniać warunki spójności, tj. uwzględniać tylko kryteria istotne z punktu widzenia problemu (takie które mają wpływ na podjęcie decyzji), wszystkie takie kryteria oraz żadnych nadmiarowych, które nic nie wnoszą do analizy problemu lub które nie pozwalają różnicować rozważanych wariantów decyzyjnych.

☀ Określenie kategorii problemu decyzyjnego

Istnieją trzy podstawowe kategorie problemów decyzyjnych:

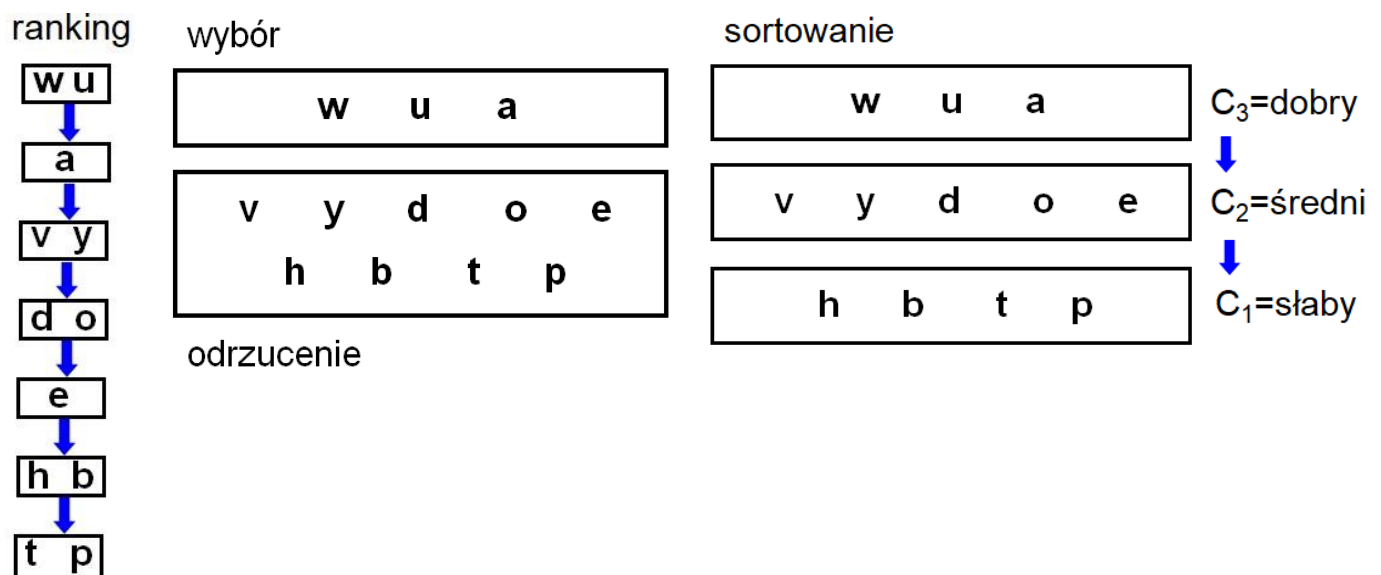
- **Porządkowanie** (tworzenie rankingu) – narzucenie na zbiór wariantów porządku, *częściowego* lub *pełnego* – *uszeregowanie wariantów od najlepszego do najgorszego*. Porządek zupełny nie dopuszcza nieporównywalności, tzn. dla wszystkich par wariantów da się określić, czy jeden jest preferowany nad drugi czy też są nierozróżnialne, tj. takie same względem problemu decyzyjnego. Porządek częściowy dopuszcza nieporównywalność.
Przykład: państwa – innowacyjność; miasta – warunki życia lub potencjał ludzki; MBA, szkoły, uniwersytety, szpitale, Webometrics.
- **Wybór** – wybór najlepszych (wyróżniających się, najbardziej preferowanych) wariantów, czyli wybór podzbioru A' zbioru A (w szczególnym przypadku A' może zawierać tylko jeden wariant).

Przykład: kupno domu lub samochodu; lokalizacja elektrowni; przebieg drogi; stypendium; kandydaci na stanowisko; system biletowy; projekt windy.

- **Sortowanie** lub klasyfikacja – podział zbioru wariantów na predefiniowane (określone z góry, przed analizą) klasy (kategorie), które w sortowaniu są uporządkowane pod względem preferencji (tj. da się powiedzieć, że jedna klasa jest bardziej preferowana od innej, a nie tylko że klasy te są różne).

Przykład: zdolność kredytowa; kwalifikacje pracowników; ryzyko; turystyka; medycyna (diagnozy, sposób leczenia), systemy uprawy; transport niebezpiecznych materiałów.

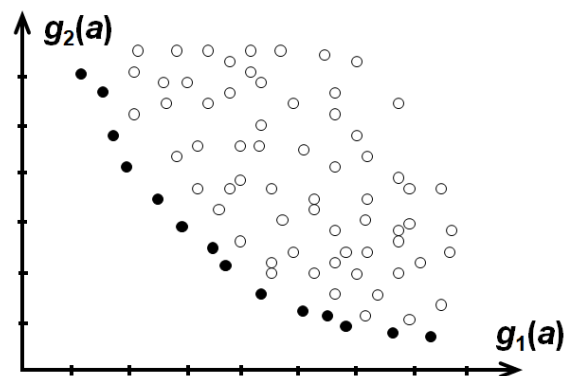
Ilustracja wyników spodziewanych w trzech rodzajach problemów decyzyjnych:



☀ Dominacja, Pareto-optymalność

- Relacja dominacji Δ wyraża jednogłośność kryteriów wobec porównania pary wariantów, dla wariantów $a, b \in A$, a dominuje b ($a \Delta b$) jeśli $g_j(a) \geq g_j(b)$ dla każdego kryterium $j \in G$
- Δ zdefiniowana na zbiorze A ma charakter porządku częściowego (jest to relacja asymetryczna i przechodnia)
- Δ jest w praktyce zwykle bardzo rzadka (np. w większości problemów za lepszą jakość płaci się większe pieniądze)
- $a \in A$ jest **Pareto-optymalny** jeśli nie istnieje $b \in A$, taki że $b \Delta a$ (b dominuje a).
- a jest **niezdominowany w sensie silnym** jeśli nie istnieje $b \in A$, który miałby oceny co najmniej tak samo dobre na wszystkich kryteriach i ściśle lepszą ocenę na co najmniej jednym kryterium
- a jest **niezdominowany w sensie słabym** jeśli nie istnieje $b \in A$, który miałby oceny ściśle lepsze na wszystkich kryteriach

Ilustracja wariantów Pareto-optimalnych dla problemu dwukryterialnego z dwoma kryteriami typu koszt:



☀ Zrób zadanie I z arkusza ćwiczeń.

Omówienie celu wszystkich zadań znajduje się na końcu tego pliku, a rozwiązane ćwiczeń - dla sprawdzenia - w dedykowanym pliku.

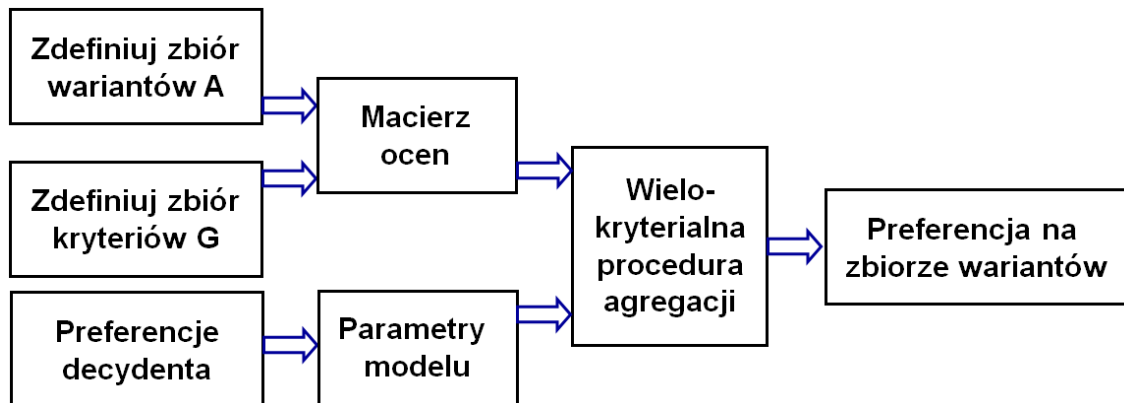
☀ Informacja preferencyjna

- Relacja dominacji Δ zostawia wiele par wariantów nieporównywalnych
- Wspomaganie decyzji wymaga wzbogacenia Δ przez dodatkową informację zwaną informacją preferencyjną
- Informacja preferencyjna odnosi się do opinii decydenta, jego systemu wartości, przekonań... dotyczących problemu decyzyjnego

Standardowo wyróżnia się informację preferencyjną wewnątrz kryterialną (ang. *intracriterion*) i międzykryterialną (ang. *intercriteria*). Ta pierwsza odwołuje się tylko do pojedynczego kryterium (np. na kryterium g_1 różnica ocena równa 5 jest zaniedbywalna, a różnica 10 jest już bardzo istotna). Zasięg tej drugiej obejmuje kilka lub wszystkie kryteria, bo podane preferencje można interpretować tylko w takim kontekście (np. sama informacja, że waga

kryterium jest równa 2 nic nam nie mówi, o ile nie znamy wag innych kryteriów;
stwierdzenie, że wariant a jest preferowany nad wariant b obejmuje wszystkie kryteria).

☀ **Wielokryterialna procedura agregacji** (ang. *MCAP – multiple criteria aggregation procedure*) - procedura, która ustala całkowitą preferencję na zbiorze wariantów A, biorąc pod uwagę informację preferencyjną oraz macierz ocen. Jest to serce metod wspomagania decyzji.



3. *Elicytacja informacji preferencyjnej*

- Zbiór wariantów $A = \{a_1, a_2, \dots, a_m\}$, zbiór kryteriów $G = \{g_1, g_2, \dots, g_n\}$ oraz relacja dominacji Δ
- Informacja preferencyjna I = każda informacja, która pozwala porównywać pary wariantów nieporównywalne względem relacji dominacji Δ , tj. uczynić warianty bardziej porównywalnymi niż wynika to tylko z ich ocen
- Proces elicytacji preferencji jest istotną składową procesu decyzyjnego:
Proces elicytacji preferencji \subset Proces wspomagania decyzji \subset Proces decyzyjny
- $I = I^{\text{in}} \cup I^{\text{res}}$ (Informacja preferencyjna może być zorientowana na wejście lub na wyjście)
- I^{in} : informacja preferencyjna zorientowana na wejście (ang. *input oriented*) odnosi się do parametrów modelu wykorzystywanego przez metodą wspomagania decyzji, np.:
 - „kryterium g_3 jest najważniejsze”
 - „miara substytucji (ang. *substitution rate*) między kryteriami g_1 oraz g_4 to 3”
 - „waga w_3 powinna wynosić co najmniej 0.4”
- I^{res} : informacja preferencyjna zorientowana na wyjście (ang. *output oriented*) odnosi do fragmentu końcowego wyniku, które chcemy uzyskać, np.:
 - „preferuję a_2 nad a_7 ”
 - „preferuję a_2 nad a_7 w większym stopniu niż preferuję a_5 nad a_1 ”
 - „ a_3 jest wariantem słabym, zaś a_6 bardzo dobrym”
 - „ a_{11} powinno zostać przypisane do klasy niegorszej niż C_3 ”
- Wartości parametrów są bez znaczenia dopóki nie określono wielokryterialnego procedury agregacji (MCAP), w ramach której będą wykorzystane

☀ **Zrób zadanie II z arkusza ćwiczeń.**

☀ Modele preferencji

Model preferencji to narzędzie matematyczne, które pozwala na reprezentację preferencji decydenta wewnątrz metody wspomagania decyzji oraz agregację ocen wariantów na poszczególnych kryteriach:

- Dzięki agregacji, warianty są bardziej porównywalne;
- Eksploatacja relacji preferencji określonej na zbiorze wariantów prowadzi do ostatecznej rekomendacji (wyboru, klasyfikacji, rankingu).

☀ Rodziny modeli preferencji:

Istnieją trzy podstawowe rodziny modeli preferencyjnych: funkcyjny, relacyjny oraz regułowy:

- **funkcja** – np. addytywna funkcja użyteczności (UTA); dla każdego wariantu wypracowujemy jedną liczbę (użyteczność, wartość), która agreguje jego oceny na wszystkich kryteriach zgodnie z preferencjami decydenta;
- **system relacyjny** – np. relacja przewyższania aSb – a jest „co najmniej tak dobry jak” b (rodzina metod ELECTRE); warianty lub obiekty są porównywane parami, ustalana jest dla nich prawdziwość lub stopień relacji, a następnie ta relacja jest eksploatowana w celu wypracowania rekomendacji;
- **zbiór reguł decyzyjnych** – wykorzystywany np. dla celów klasyfikacji w ramach teorii zbiorów przybliżonych; reguły mają postać logicznych stwierdzeń typu „*jeżeli* są spełnione pewne warunki, *to* zachodzi określona decyzja”; model regułowy jest najbardziej zrozumiały dla ludzi oraz najbardziej ogólny ze wszystkich trzech rodziny modeli:

jeżeli $g_1(a) > 10$ oraz $g_2(a) < 20$ **to** klasa C2 lub wyższa

jeżeli $d_1(a,b) > 2$ oraz $d_2(a,b) > -1$ **to** aSb (gdzie $d_1(a,b) = g_1(a) - g_1(b)$, a S jest relacją przewyższania)

☀ Modelowanie preferencji w ramach procedury P

Modelowanie preferencji polega na przetłumaczeniu informacji preferencyjnej w kategoriach wartości parametrów założonego modelu preferencji P , który odpowiada procedurze wykorzystywanej następnie do agregacji ocen wariantów.

Procedura P , w której do określenia jest k parametrów preferencji $v = (v_1, v_2, \dots, v_k)$.

Przykład: P = suma ważona $WS(a) = \sum_{i=1, \dots, k} w_i \cdot g_i(a)$; w = wagi = (w_1, w_2, \dots, w_k) .

Ω - przestrzeń akceptowalnych wartości v w przypadku, gdy brak informacji preferencyjnej.

Przykład: $\Omega = \{w_i \geq 0, i \in G : \sum_{i \in G} w_i = 1\}$.

Wiedza dotycząca v , która wynika z informacji preferencyjnej I jest zdefiniowana jako zbiór ograniczeń dotyczących v oznaczany jako $\Omega(I) \subseteq \Omega$.

Przykład: Jeżeli $I = \{(11, 10, 10) > (10, 12, 10)\}$, to $\Omega(I) = \{w \in \Omega : w_1 > 2w_2\}$, gdyż:

$11w_1 + 10w_2 + 10w_3 > 10w_1 + 12w_2 + 10w_3$ prowadzi do $w_1 > 2w_2$.

W szczególnym wypadku: $\Omega(I) = \{w\}$ – wartość każdego parametru jest ściśle określona

Przykład: szczególny przypadek: $\Omega(I) = \{w\}$, np. gdy $(10,10,11) \sim (10,11,10) \sim (11,10,10)$.

W przeciwnym razie, wartość co najmniej jednego parametrów jest nieprecyzyjna.

Jeśli określona jest procedura MCAP oznaczona symbolem P , proces elicytacji informacji preferencyjnej polega na interakcji między decydentem oraz analitykiem, która prowadzi decydenta do udzielenia informacji na temat swoich preferencji w ramach P . *Najczęściej odbywa się to z wykorzystaniem sekwencji Q/A (pytań/odpowiedzi), w której decydent przyrostowo wyraża swoje preferencje. Taka informacja preferencyjna jest tłumaczona do zbioru $\Omega(I) \subseteq \Omega$ dozwolonych wartości parametrów procedury P . Na końcu procesu, wykorzystanie $\Omega(I)$ powinno prowadzić, poprzez wykorzystanie P , do wyniku, który jest spójny z preferencjami decydenta, tj. do wypracowania rekomendacji.*

4. Natura procesu elicytacji informacji preferencyjnej

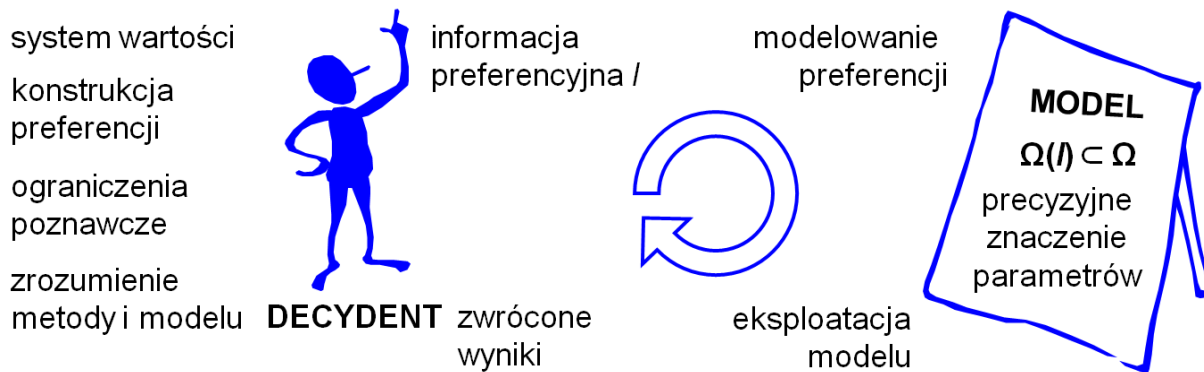
Proces elicytacji informacji preferencyjnej może mieć charakter deskrypcyjny lub konstruktywistyczny w zależności od typu decydenta/użytkownika, z którymi jako analitycy mamy do czynienia. Zdecydowanie łatwiejsza i bardziej przyjemna jest praca w ramach podejścia konstruktywistycznego.

Podejście deskrypcyjne (opisowe, ang. *descriptivist*)

- Sposób, w jaki warianty powinny być ze sobą porównane jest zdefiniowany w głowie decydenta przed rozpoczęciem procesu elicytacji
- Proces elicytacji nie zmienia istniejącej struktury preferencji
- Informacja preferencyjna jest stabilna
- Model preferencji powinien reprezentować istniejące preferencje tak wiarygodnie i wiernie, jak to tylko możliwe

Podejście konstruktywistyczne (ang. *constructivist*)

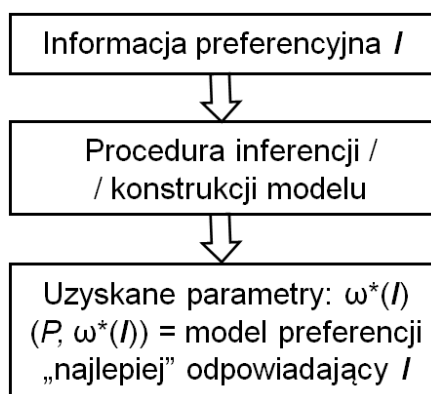
- Preferencje decydenta nie są w pełni ugruntowane
- Celem procesu elicytacji preferencji jest określenie lub nawet zmienienie istniejących elementów preferencji
- Proces elicytacji ma dać konkretny obraz przeświadczenia decydenta na temat sposobu porównania wariantów
- W podejściu tym szczególnie istotne jest uwzględnienie, że decydent, z którym pracujemy lub którym sami jesteśmy ma swój system wartości/preferencji, ale też to, co nam powie zależy od chwili, w której o to prosimy (preferencje są konstruowane, a nie tylko ujawniane), ma on swoje ograniczenia tym poznawcze i dotyczące zrozumienia metody lub modelu, w ramach którego jego preferencje będą wykorzystywane.



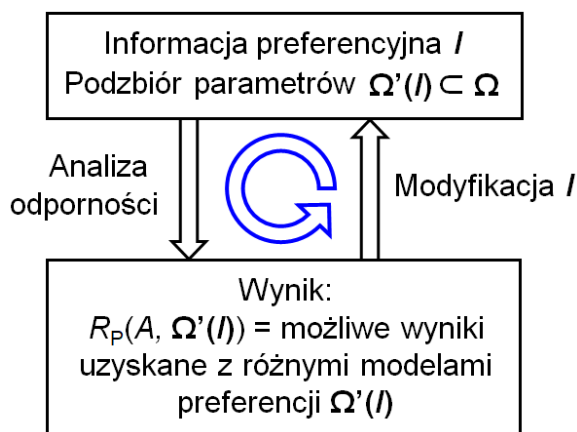
5. Narzędzia elicytacji informacji preferencyjnej

Wyróżniamy trzy podstawowe narzędzia elicytacji informacji preferencyjnej w zależności od spójności tej informacji (spójna z założonym modelem preferencji lub nie) oraz sposoby wykorzystania modelu preferencji, którego wartości parametrów są określone na podstawie tej informacji preferencyjnej.

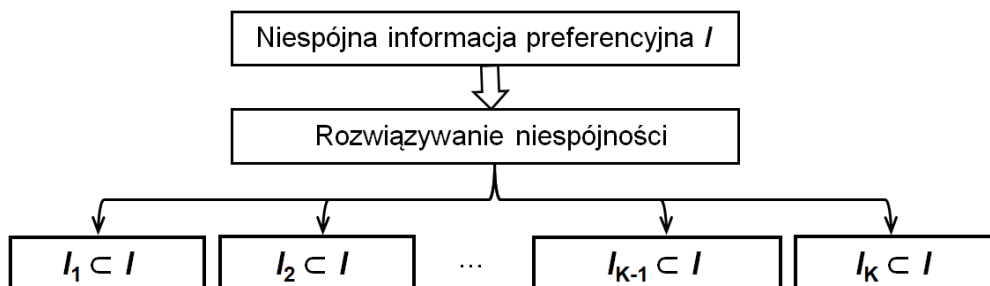
- **Agregacja / dezagregacja:** procedura inferencji parametrów na podstawie informacji preferencyjnej zorientowanej na wyjście; ustalany jest jeden model preferencji (a ściślej rzecz biorąc - jego instancją z konkretnymi wartościami parametrów), który najlepiej odpowiada przedstawionym preferencjom.



- **Analiza odporności rozwiązań:** na podstawie informacji preferencyjnej pośredniej, tj. zorientowanej na wyjście, ograniczany jest zbiór instancji modelu preferencji spójnych z preferencjami decydenta; zamiast wyboru jednej takiej instancji, wykorzystywane są one wszystkie do badania odporności rozwiązań, tj. stabilności wyników w zależności od wykorzystanej spójnej instancji modelu preferencji; wyniki takiej analizy można wykorzystać do stymulowania decydenta do modyfikacji (zmiany lub wzbogacenia) swojej informacji preferencyjnej, która dalej zawęży zbiór spójnych instancji, co z kolei prowadzi do bardziej precyzyjnego wyniku.



- **Wykrycie niespójności w informacji preferencyjnej i jej rozwiązanie:** w przypadku, gdy informacja preferencyjna jest niespójna nie istnieje żadna instancja modelu preferencji, która ją odtwarza. W takim scenariuszu należy pomóc decydentowi w przywróceniu spójności. Rozwiązanie niespójności polega na wskazaniu minimalnych podzbiorów elementów informacji preferencyjnej I_1, \dots, I_K , które należy usunąć lub zmienić, aby informacja stała się spójna.



6. Przypomnienie lub omówienie metod poznanych na Wspomaganiu Decyzji poprzez rozwiązanie zadań ilustrujących

☀ **Zrób zadania III-XV z arkusza ćwiczeń.**

W razie potrzeby, skonsultuj cel oraz omówienie poszczególnych zadań z opisem znajdującym się w następnym rozdziale.

7. Opis celu zadań ilustrujących z arkusza ćwiczeń

Zad. I. Umiejętność praktycznej identyfikacji wariantów niezdominowanych w sensie słabym i silnym

- Interpretacja kryteriów o różnych kierunkach preferencji (g_1 - typu koszt (im niższa ocena, tym lepsza), g_2 - typu zysk (im wyższa ocena, tym lepsza))
- Wystawianie stożków dominacji dla danego wariantu z tworzącymi w kierunkach preferencji poszczególnych kryteriów (patrz wariant c) - np. dla g_1 w lewo (w stronę lepszych, tj. mniejszych, ocen), dla g_2 w górę

- Dla wariantu niezdominowanego w sensie silnym nie może być innego wariantu wewnątrz stożka ani na jego tworzących (np. dla c istnieje a ; dla b istnieje h ; dla a oraz h nie istnieje nic)
- Dla wariantu niezdominowanego w sensie słabym nie może być innego wariantu wewnątrz stożka (dla c istnieje a , dla n nie istnieje nic)
- Przy większej liczbie wymiarów (kryteriów) należy odwołać się bezpośrednio do definicji, tj. dla danego wariantu udowodnić, że nie istnieje inny, który spełniałby określone warunki
- Wariant niezdominowany w sensie silnym jest też niezdominowany w sensie słabym

Zad. II. Wartości parametrów modelu preferencji (np. wag) nabierają znaczenia dopiero w kontekście procedury (metody) wielokryterialnej, w której będą wykorzystane

- Przy tych samych wartościach parametrów, interpretowanych w inny sposób w różnych metodach lub modelach, wyniki mogą się znacząco różnić
- Suma ważona = suma iloczynów wag i ocen; agregacja typu Condorceta = suma wag skojarzonych z kryteriami, dla których jeden wariant jest co najmniej tak dobry jak inny wariant
- Suma ważona - a preferowany nad b , agregacja typu Condorceta - b preferowany nad a
- Preferencja jest odwrotna przy tych samych wartościach ocen oraz wag, przy czym wagi te są interpretowane i wykorzystywane w inny sposób w dwóch przedstawionych modelach

Zad. III. Przykład wielokryterialny dotyczący wyboru lokalizacji elektrowni

- Sześć wariantów ocenionych na trzech kryteriach
- Różne kierunki preferencji kryteriów: typu zysk i koszt
- Kolorem ciemnym szarym wyróżniono najlepsze oceny na poszczególnych kryteriach, jasnym szarym - najgorsze (zwróć uwagę na kierunki preferencji)
- Przykład będzie wykorzystywany dla ilustracji zasady działania różnych metod w ramach kolejnych zajęć

Zad. IV. Metoda UTA

Omówiona szczegółowo w ramach przedmiotu *Wspomaganie Decyzji (WD) - lab 7* - służy do wspomagania problemu **porządkowania (rankingu)** - uzyskany ranking jest zupełny (nie ma w nim nieporównywalności, tj. dla każdej pary wariantów da się powiedzieć, czy jeden jest lepszy (preferowany) od drugiego czy też są takie same (nierozróżnialne))

Metoda UTA wykorzystuje model preferencji w postaci **addytywnej funkcji użyteczności**:

- Z oceną każdego wariantu skojarzona jest **wartość (użyteczność) cząstkowa**, którą należy odczytać z wykresu - jest ona miarą jakości przypisaną konkretnej ocenie z punktu widzenia decydenta
- Dla danego wariantu wartości cząstkowe odpowiadające jego ocenom na poszczególnym kryteriach sumuje się, by uzyskać **wartość całkowitą (globalną)**

- We wszystkich metodach opartych na teorii wartości (użyteczności) podstawowy pomysł zakłada agregację ocen wariantu do **jednej liczby**, która odzwierciedlałaby jego całkowitą jakość (jak dobry/zły jest z punktu widzenia wszystkich kryteriów). Jeśli z każdym wariantem skojarzona jest taka liczba, to możemy na zbiór wariantów narzucić porządek zupełny, tj. utworzyć ranking, uporządkowując je od najlepszego do najgorszego.
- **Funkcje wartości cząstkowych** w metodzie UTA mają charakter **liniowych odcinkami**, tzn. można zdefiniować liczbę punktów charakterystycznych (złamań) dla danej funkcji; Im więcej takich punktów, tym większa swoboda w dopasowaniu do preferencji użytkownika; funkcje te muszą mieć **charakter monotoniczny** zgodny z kierunkiem preferencji danego kryterium (dla kryterium typu zysk - niemalejący; dla kryterium typu koszt - nierosnący)
- Wartość całkowita (globalna) **normalizowana jest do przedziału od 0 do 1**

Metoda UTA wykorzystuje informację preferencyjną w postaci **rankingu wariantów referencyjnych**:

- Warianty referencyjne stanowią zazwyczaj mały podzbiór całego zbioru wariantu - najczęściej są to warianty najlepiej znane decydentowi lub takie, dla których najłatwiej przedstawić mu przykładową decyzję

"Silnikiem metody UTA" jest sformułowanie i rozwiązanie **problemu regresji porządkowej**, który sprawdza, czy istnieje taka funkcja wartości, która przy zachowaniu warunków monotoniczności i normalizacji odtworzy podany przez decydenta ranking referencyjny:

- W przypadku znalezienia takiej funkcji, możliwe jest odczytanie wartości całkowitej dla każdego wariantu (także tych niereferencyjnych), a w konsekwencji uzyskanie rankingu zupełnego dla całego zbioru wariantów (jeśli wyjąć z tego rankingu tylko warianty referencyjne, to będzie dla nich odtworzony porządek podany przez decydenta)
- Formalnie podejście takie reprezentuje paradygmat **dezagregacji/agregacji** - z małego pożądanego fragmentu końcowego wyniku (=ranking dla wariantów referencyjnych) wyprowadzany jest model preferencji (*dezagregacja*), który następnie stosuje się do oceny całego zbioru wariantów (*agregacja*)
- W ogólności **istnieje wiele funkcji wartości spójnych z preferencjami decydenta**, UTA zakłada wybór jednej z nich (przy pomocy domyślnej reguły lub interakcji z użytkownikiem - patrz program Visual UTA)

Zad. V. Zapis problemu regresji porządkowej w metodzie UTA

- Regresja porządkowa jest problemem **programowania liniowego**
- Sformułowanie problemu uwzględnia trzy typy ograniczeń:
odtworzenie rankingu referencyjnego (rozbicie wartości całkowitej wariantów referencyjnych na sumę użyteczności cząstkowych odpowiadających im ocenom), np. skoro

AUT ma być preferowana na BEL, to użyteczność AUT, $U(AUT)$, musi być wyższa niż użyteczność BEL, $U(BEL)$; w przypadku nierozróżnialności - użyteczności musiałyby być równe;

zapewnienie normalizacji (użyteczność najgorszych ocen na poszczególnych kryteriach musi być zerowa, suma użyteczności najlepszych ocen na wszystkich kryteriach musi dawać jeden),

zapewnienie monotoniczności funkcji wartości częściowych poprzez porównanie wartości częściowych w kolejnych punktach charakterystycznych (uwaga na kierunek preferencji danego kryterium), tj. w lepszym punkcie charakterystycznym użyteczność musi być co najmniej tak dobra jak w punkcie charakterystycznym gorszym.

- Przy zamodelowanie relacji preferencji, najłatwiej **przekształcić nierówność ostrą ($>$) w nieostrą (\geq) dodając wartość epsilon** (w domyśle - arbitralnie wybranej bardzo małej liczby)
- Jeśli potraktować epsilon jego zmienną i maksymalizować jego wartość w funkcji celu, to funkcja wartości spójna z preferencjami decydenta istnieje, jeżeli układ ograniczeń jest niesprzeczny a optymalna (największa) wartość epsilon jest większa od zera
- W układzie ograniczeń dodatkowo należy zamodelować wartości częściowe odpowiadające ocenom wypadającym pomiędzy punktami charakterystycznymi;

Dla problemu regresji porządkowej zmiennymi są tylko wartości częściowe w punktach charakterystycznych; wszystkie pozostałe należy do nich odnieść z wykorzystaniem interpolacji liniowej (wartość w poprzednim punkcie charakterystycznym + wartość wynikająca z zastosowania twierdzenia Talesa)

Przedstawiono to na przykładzie użyteczności dla oceny 400, wypadającej pomiędzy punktami charakterystycznymi w ocenach 200 oraz 600; startujemy z użyteczności punktu gorszego (tu 600), do której dodawana jest odpowiednia proporcja różnicy użyteczności między sąsiednimi punktami charakterystycznymi (tu 200 i 600)

Zad. VI. Wybór przykładowej funkcji wartości spójnej z rankingiem podanym przez decydenta, a przy tym zachowującej monotoniczność i spełniającej normalizację

- Narysowano częściowe funkcje wartości na dwóch pierwszych kryteriach (są liniowe, bo miały być tylko dwa punkty charakterystyczne, czyli punkty w których może nastąpić "złamanie funkcji"), jeśli odczytać wartości sumaryczne dla wariantów referencyjnych, mamy:
 $U(AUT)=0.12+0.32=0.44 > U(ITA)=0.24+0.16=0.4 > U(BEL) = 0.0+0.0=0.0$,
czyli porządek jest zgodny z rankingiem który mieliśmy odtworzyć ($AUT > ITA > BEL$)
- Dorysowanie funkcji wartości częściowych dla trzeciego kryterium musi gwarantować:

normalizację (najlepszej ocenie (200) musi odpowiadać 1 - 0.3 (max na g_1) - 0.4 (max na g_2) = 0.3, a najgorszej (1000) - 0.0);

monotoniczność - wartość w punkcie 200 musi być większa lub równa od wartości w punkcie 600; z kolei wartość w tym punkcie musi być większa równa od wartości w punkcie 1000;

odtworzenie rankingu referencyjnego (AUT oraz ITA mają takie same oceny (600), więc w tym wypadku nic nie zepsujemy, jakkolwiek prowadząc funkcję);

- Dla przykładu zakładamy funkcję liniową (choć dozwolone jest dla tej funkcji jedno złamanie w środkowym punkcie charakterystycznym = 600); ostatecznie:

$$U(\text{AUT}) = 0.12 + 0.32 + 0.15 = 0.59 > U(\text{ITA}) = 0.24 + 0.16 + 0.15 = 0.55 >$$

$$U(\text{BEL}) = 0.0 + 0.0 + 0.3 = 0.3,$$

czyli ranking referencyjny został odtworzony ($\text{AUT} > \text{ITA} > \text{BEL}$)

- Z funkcji wartości można odczytać wartości cząstkowe, a potem całkowite dla wszystkich sześciu wariantów - ranking uzyskany dzięki temu plasuje FRA na pierwszej pozycji, a BEL na ostatniej (jeśli wyjąć z niego warianty referencyjne - porządek podany przez decydenta jest zachowany)

Zad. VI (kontynuacja) - ilustracja istnienia wielu funkcji wartości spójnych z preferencjami decydenta

- Jeśli wykres funkcji wartości cząstkowej na trzecim kryterium narysować inaczej (np. najpierw na poziomie 0.3 od 200 do 600, a potem liniowo w dół od 600 (0.3) do 1000 (0.0)), to ranking referencyjny wciąż jest odtworzony ($\text{AUT} > \text{ITA} > \text{BEL}$), ale ranking dla wszystkich pozostałych wariantów może być inny (patrz pierwsza pozycja (FRA vs. AUT) lub odwrócona relacja preferencji między ITA i FRA)
- **Temat ten jest przedmiotem laboratorium dotyczącego Odpornej Regresji Porządkowej**

Zad. VII. Metody ELECTRE

Rodzina metod ELECTRE jest oparta na **relacyjnym modelu preferencji** - w pierwszym etapie konstruuje one tzw. relację przewyższania, a w drugim - eksploatują ją tak, by wypracować rekomendację w duchu problemu wyboru rankingu lub sortowania

Metoda ELECTRE Is (czytaj: "jeden es") jest omówiona szczegółowo w ramach przedmiotu *Wspomaganie Decyzji (WD)* - lab 9 - służy do wspomagania problemu **wyboru** (wskazuje podzbiór najlepszych, tj. najbardziej preferowanych, wariantów)

Relacyjny model preferencji:

- Intuicyjna interpretacja **relacji przewyższania S** to "**co najmniej tak dobry jak**"

- Relacja przewyższania grupuje w sobie relację silnej preferencji, słabej preferencji oraz nierozróżnialności (jeśli a przewyższa b , to a jest lepszy (preferowany) lub taki sam (nierozróżnialny) jak b)
- Sprawdzając prawdziwość relacji przewyższania dla pary obiektów (w szczególności wariantów), można wnioskować o bardziej szczegółowej relacji, która między nimi zachodzi:
jeśli aSb i bSa , to aIb (nierozróżnialność),
jeśli aSb i $\text{not}(bSa)$ to $a>b$ (preferencja),
jeśli $\text{not}(aSb)$ i $\text{not}(bSa)$, to $a?b$ (nieporównywalność) - w teorii wartości (użyteczności) nie da się zamodelować nieporównywalności

Weryfikacja prawdziwości relacji przewyższania opiera się na realizacji dwóch **testów, zgodności i niezgodności**:

- Test zgodności bada, czy **siła koalicji kryteriów, które wspierają relację przewyższania jest wystarczająco wysoka**, co jest warunkiem koniecznym przewyższania
- Test niezgodności bada, czy istnieje **co najmniej jedno takie kryterium (lub w ogólności koalicja kryteriów), które wskazywałyby że jeden wariant jest krytycznie gorszy od drugiego**, co wykluczałoby przewyższanie

Metody ELECTRE wymagają podania szeregu parametrów o intuicyjnej interpretacji:

- **Próg nierozróżnialności** q_i na każdym kryterium wskazuje, jaka różnica w ocenach jest wg decydenta **zaniedbywalna** (tzn. jeden wariant może być gorszy od drugiego o co najwyżej próg nierozróżnialności, a wciąż jest postrzegany jako nierozróżnialny z nim)
- **Próg preferencji** p_i na każdym kryterium wskazuje, jaka różnica w ocenach jest wg decydenta **istotna** (tzn. jeden wariant lepszy od drugiego o co najmniej próg preferencji jest postrzegany jako silnie preferowany na tym kryterium; jeśli przewaga jednego wariantu nad drugim jest większa od progu nierozróżnialności a mniejsza od progu preferencji, to mówimy wtedy o słabej preferencji)
- **Próg veta** v_i na każdym kryterium wskazuje, jaka różnica w ocenach jest wg decydenta **krytyczna** (tzn. jeden wariant gorszy od drugiego o co najmniej próg veta jest postrzegany jako krytycznie gorszy na tym kryterium; jest to słabość na tyle duża, że nawet jeśli byłby lepszy na wszystkich pozostałych kryteriach, to i tak nie może zająć relacja przewyższania, czyli nie można powiedzieć na poziomie globalnym, wariant ten jest co najmniej tak samo dobry jak inny wariant)
- **Waga** k_i dla każdego kryterium wskazuje na istotność danego kryterium względem pozostałych; wagi w metodach ELECTRE mają interpretację czerpiącą z teorii społecznego wyboru - **siła głosu danego kryterium**
- **Próg odcięcia** λ , który wskazuje jaka powinna być **minimalna siła** koalicji kryteriów wspierających relację przewyższania (w ELECTRE Is) lub jaka powinna być minimalna

wiarygodność relacji przewyższania (w ELECTRE TRI), **aby mieć podstawy do uznania relacji przewyższania za prawdziwą**

Zad. VIII. Weryfikacja prawdziwości relacji przewyższania w metodzie ELECTRE Is:

- Prawdziwość relacji przewyższania sprawdza się dla każdej pary wariantów (tzn. porównuje się każdy wariant z wszystkimi pozostałymi)
- Przeprowadzenie testu zgodności opiera się na obliczeniu **globalnego współczynnika zgodności** - jest on średnią ważoną cząstkowych współczynników zgodności (każdy z nich wyraża, czy wariant *a* przewyższa wariant *b* na skali od 0 (brak przewyższania) do 1 (przewyższanie);

Cząstkowy współczynnik zgodności uzyskuje się poprzez porównanie różnic w ocenach z progami nierozróżnialności i preferencji;

Test zgodności jest zaliczony, jeśli globalny współczynnik zgodności (odzwierciedlający siłę koalicji kryteriów wspierających relację przewyższania) jest większy lub równy od progu odcięcia (wtedy koalicja jest wystarczająco silna)

- Przeprowadzenie testu niezgodności opiera się na sprawdzeniu, czy nie istnieje przynajmniej jedno kryterium na którym wariant byłby krytycznie (o co najmniej **próg veta**) gorszy od drugiego wariantu; test niezgodności jest niezaliczony/oblany, jeśli takie kryterium istnieje
- **Relacja przewyższania jest prawdziwa** (zachodzi), jeżeli test zgodności i niezgodności są zaliczone, tj. siła koalicji kryteriów wspierających hipotezę o przewyższaniu jest wystarczająco silna i brak kryterium, na którym wariant byłby krytycznie gorszy
- **Relacja przewyższania jest fałszywa** (nie zachodzi), jeżeli test zgodności albo test niezgodności są oblany

Eksploatacja relacji przewyższania w metodzie ELECTRE Is:

- Wyniki weryfikacji prawdziwości relacji przewyższania dla każdej pary wariantów gromadzone są w macierzy **relacji przewyższania** (1 - zachodzi, 0 - nie zachodzi), którą następnie reprezentuje się jako **graf przewyższania** (warianty - wierzchołki, łuki - prawdziwość relacji przewyższania (w grafie nie rysuje się łuku od wierzchołka do samego siebie))
- Podzbiór najlepszych wierzchołków to **jądro grafu** relacji przewyższania - zbiór wierzchołków spełniających warunki stabilności wewnętrznej (wierzchołki w jądrze nie są bezpośrednio połączone łukiem) i stabilności zewnętrznej (do wierzchołków poza jądrem musi dochodzić łuk od jakiegokolwiek wierzchołka w jądrze)
- Procedura poszukiwania jądra grafu daje rozwiązanie unikalne, jeśli graf jest acykliczny (jeśli cykl istnieje, należy go usunąć - jeden ze sposobów zakłada agregację wszystkich

wierzchołków wchodzących w skład cyklu do wierzchołka złożonego, który przejmuję cechy wierzchołków składowych)

- Aby znaleźć jądro grafu, odnajduje się wierzchołki, do których nie dochodzi żaden łuk (muszą one znaleźć się w jądrze ze względu na stabilność zewnętrzną), a potem podejmuje się decyzji dla tych wierzchołków, o których wszystkich poprzednikach wiadomo, czy są w jądrze czy poza nim.

Zad. IX. Sprawdzenie prawdziwości relacji przewyższania dla przykładowej pary wariantów

- Sposób rozumowania dla pary $(SWE, AUT) = (S, A)$:
 g_1 : S jest lepsze od A o 8 - niezależnie od wartości progów zgodność = 1 (skoro S jest lepsze od A, to jest co najmniej tak samo dobre), brak niezgodności (skoro jest jakakolwiek zgodna, tu nawet całkowita, to nie może być niezgodności z hipotezą o przewyższaniu),
 g_2 : S jest lepsze od A o 2 - niezależnie od wartości progów zgodność = 1, brak niezgodności,
 g_3 : S jest gorsze od A o 400 - 400 to więcej niż próg preferencji = 200, czyli zgodność = 0 (S jest gorsze znacząco, o co najmniej próg preferencji), ale mniej niż próg veta = 700, czyli niezgodności też nie ma (bo ta dopiero by wystąpiła, jeśli S byłoby gorsze o co najmniej próg veta);
całkowity współczynnik zgodności = 0.5 (średnia ważona współczynników cząstkowych = $(2*1+3*1+5*0)/(2+3+5)$) jest mniejszy od progu odcięcia = 0.8;
W konsekwencji test zgodności jest niezaliczony i relacja przewyższania nie zachodzi: $\text{not}(SWE \succ S \succ AUT)$;
• Sposób rozumowania dla pary (A, S) : całkowity współczynnik zgodności = 0.85 jest większy lub równy od progu odcięcia = 0.8 (np. na g_1 S jest gorsze od A o 8, ale 8 to dokładnie próg nierozróżnialności, czyli całkowita zgodność, na g_2 S jest gorsze od A o 2, a to w połowie między progiem nierozróżnialności (1) i preferencji (3), czyli zgodność = 0.5);
Brak veta na jakimkolwiek kryterium;
W konsekwencji relacja przewyższania zachodzi: $AUT \succ S \succ SWE$.

Zad. X. Uzupełnienie macierzy relacji przewyższania na podstawie wyników testów zgodności i niezgodności oraz znalezienie jądra grafu acyklicznego

- Jeżeli test zgodności jest zaliczony (dla naszego przykładu - całkowity współczynnik zgodności ≥ 0.8) i brak veta na jakimkolwiek kryterium (F; przy sprawdzeniu niezgodności można ograniczyć się tylko dla par, dla których test zgodności jest zaliczony, w przeciwnym razie i tak relacja przewyższania jest fałszywa, więc nie ma to sensu), to relacja przewyższania zachodzi (1); w przeciwnym razie nie (0)

- **Graf** relacji przewyższania: wierzchołki = warianty, łuki = prawdziwość relacji przewyższania (nie reprezentuje się zwrotność relacji przewyższania, czyli "1" na głównej przekątnej, które byłyby reprezentowane jako łuki od wierzchołków do samego siebie)
- **Iteracyjna procedura identyfikacji jądra - ilustracja:**
Wierzchołki, do których nie dochodzi żaden łuk (G, B) do jądra - stabilność zewnętrzna;
Szukamy wierzchołków, o których wszystkich poprzednikach wiemy, czy należą do jądra czy też nie - wierzchołki, do których dochodzi łuk od G (A,F,I) poza jądro - stabilność wewnętrzna;
Szukamy wierzchołków, o których wszystkich poprzednikach wiemy, czy należą do jądra czy też nie - wierzchołek S do jądra - stabilność zewnętrzna
Kernel (jądro) = {G,B,S} - wskazany jako podzbiór najlepszych wariantów, pozostałe warianty (A,F,I) są odrzucone

Zad. XI. Metoda ELECTRE TRI

Metoda ELECTRE TRI jest omówiona szczegółowo w ramach przedmiotu *Wspomaganie Decyzji (WD) - lab 10* - służy do wspomagania problemu **sortowania** (przydziela warianty do predefiniowanych (określonych z góry) i uporządkowanych klas (tak, że wiadomo, że warianty, które trafią do klasy lepsze będą uznane za bardziej preferowane od wariantów, które trafią do klas gorszej))

Warianty porównuje się ze zdefiniowanymi przez decydenta **profilami (sztucznymi wariantami) separującymi klasy**. Wynikająca z tego porównania relacja przewyższania jest następnie eksploatowana z wykorzystaniem procedury pesymistycznej lub optymistycznej w celu przydziału wariantów do klas decyzyjnych.

Zad. XII. Podstawowe różnice między metodami ELECTRE TRI oraz ELECTRE Is:

- **Warianty porównuje się z profilami separującymi** (które stanowią element informacji preferencyjnej decydenta), a nie ze sobą; profil separujący stanowi umowną górną granicę gorszej klasy, jednocześnie będąc dolną granicą klasy lepszej; istotne są profile oddzielające klasy - można więc pominąć w rozważaniach profil dolny klasy najgorszej i profil górny klasy najlepszej (one i tak na nic nie wpływają)
- **Progi nierozróżnialności, preferencji i veta mają charakter stałych** - podaje się je dla każdego profilu na każdym kryterium; w ELECTRE Is progi mogą mieć charakter funkcji afinicznej, np. $q_i(a) = \alpha \cdot g_i(a) + \beta$, która uzależnia wartość progu od oceny wariantu
- **Test niezgodności na każdym kryterium może dać wynik rozmyty** - między 0 i 1, podczas gdy w ELECTRE Is była to tylko 1 (veto) lub 0 (brak veta); cząstkowy współczynnik niezgodności uzyskuje się porównując różnicę ocen z progami veta i preferencji

- **Wyniki testów zgodności i niezgodności agreguje się do jednej miary - wiarygodności relacji przewyższania**, którą dopiero zestawia się z progiem odcięcia w celu stwierdzenia, czy wiarygodność jest wystarczająco wysoka, by relacja przewyższania zachodziła; Punktem wyjścia do obliczenia wiarygodności jest całkowity współczynnik zgodności, który może zostać obniżony przez odpowiednio wysokie cząstkowe współczynniki niezgodności (patrz warunek wejścia za znak mnożenia)

Sprawdzenie prawdziwości relacji przewyższania dla przykładowych par wariantów i profili:

- Sposób rozumowania dla pary (ITA, B2):
 g_1 : ITA jest lepszy od B2 o 5 - zgodność = 1 (bo lepszy, więc co najmniej tak samo dobry), niezgodność = 0;
 g_2 : ITA jest gorszy od B2 o 4 - progi dla badania zgodności $q_2(B2) = 0$, $p_2(B2) = 1$ (zwróć uwagę, aby odwołać się do progów na odpowiednim kryterium i zdefiniowanych dla odpowiedniego profilu), a więc zgodność=0 (bo gorszy o co najmniej próg preferencji), progi dla badania niezgodności $p_2(B2) = 1$, $v_2(B2)=3$, a więc niezgodność=1 (bo gorszy o co najmniej próg veta);
 g_3 : ITA gorszy od B2 o 200, progi badania zgodności $q_3(B2) = 50$, $p_3(B2) = 100$, a więc zgodność=0 (niezgodności (efektu veta) na tym kryterium nie rozważa się);
 Całkowity współczynnik zgodności = 0.5 (średnia ważona współczynników cząstkowych);
 Wiarygodność = $0.5 \cdot (1-1)/(1-0.5)=0$ (na g_2 cząstkowy współczynnik niezgodności (1) jest większy od globalnej zgodności (0.5), a więc wchodzimy za znak mnożenia);
 Wiarygodność = 0 jest mniejsza od progu odcięcia = 0.8, a więc relacja przewyższania nie zachodzi not(ITA S B2)
- Wyniki wskazują na not(ITA S B2) oraz not(B2 S ITA), co w konsekwencji prowadzi do wniosku, że ITA oraz B2 są nieporównywalne (ITA ? B2)

Zad. XIII. Procedury przydziału do klas w metodzie ELECTRE TRI

- Procedury pesymistyczna i optymistyczna mają **charakter rozłączny**, tj. w rzeczywistych problemach powinno wykorzystywać się jedną z nich
- **Pesymistyczna** porównuje wariant z profilami, zaczynając od góry (od najlepszego) i szuka pierwszego profilu, który jest przewyższany przez wariant - szukamy więc podłogi, poniżej której nie spadniemy; rekomendowana jest klasa znajdująca się tuż nad profilem (dla ITA zatrzymujemy się na B1, więc klasa AVE; dla GER nie zatrzymujemy się na żadnym profilu, więc klasa najgorsza)
- **Optymistyczna** porównuje wariant z profilami, zaczynając od dołu (od najgorszego) i szuka pierwszego profilu, który jest preferowany nad wariant (ściśle lepszy; uwaga na różne warunki stopu w pesymistycznej i optymistycznej) - szukamy więc sufitu, powyżej którego

nie przejdziemy; rekomendowana jest klasa znajdująca się tuż pod profilem (dla BEL zatrzymujemy się na B1, więc klasa BAD, dla SWE nie zatrzymujemy się na żadnym profilu, więc klasa najlepsza)

- Klasa wynikająca z zastosowania procedury optymistycznej jest zawsze co najmniej tak dobra jak ta, wynikając z procedury pesymistycznej
- Różnice we wskazaniach dwóch procedur pojawiają się tylko, gdy wariant jest nieporównywalny (?) z co najmniej jednym profilem (patrz ITA, GER, SWE, FRA)

Zad. XIV. Zbiory przybliżone

Zbiory przybliżone są omówione w ramach przedmiotu *Wspomaganie Decyzji (WD) - lab 4* - służą do wspomagania problemu **klasyfikacji** (przydziału wariantów do predefiniowanych nieuporządkowanych klas)

- Tabela decyzyjna: zbiór wariantów opisany jest na kilku **atrybutach warunkowych** (atrybuty, w przeciwieństwie do kryteriów, nie mają określonego kierunku preferencji - np. MOC, KOSZT, BEZP.) oraz **atrybucie decyzyjnym** (np. KLASA), który mówi o przynależności danego wariantu do konkretnej klasy (zbioru)
- Potrzeba wykorzystania zbiorów przybliżonych występuje wtedy, gdy nie można zastosować zbiorów definiowalnych, tj. takich, które da się jednoznacznie zdefiniować poprzez własności elementów wchodzących w ich skład (zwróć uwagę, że warianty LUX i AUT są nierozróżnialne ze względu na wszystkie atrybuty warunkowe, a należą do różnych klas - jest to niespójność względem zgromadzonej wiedzy/opisu; zwróć uwagę, że nie należy traktować jako niespójności opisu dla NOR i FIN, które mają te same wartości na atrybutach warunkowych, ale też należą do tej samej klasy)

Zbiory przybliżone określa z wykorzystaniem pojęć:

- **Dolnego przybliżenia** zbioru - intuicyjnie są to te elementy, które na pewno należą do tego zbioru (tj. w świetle posiadanej wiedzy nic nie poddaje w wątpliwość ich przynależności)
- **Górnego przybliżenia** zbioru - intuicyjnie są to te elementy, które być może należą do tego zbioru (tj. w świetle posiadanej wiedzy nie można ich przynależności wykluczyć)
- Dla klas BAD oraz AVE istnieje różnica między górnym a dolnym przybliżeniem (formalnie nazywana **brzegiem zbioru** (klasy)), gdyż istnieje niespójność co do przynależności do klas dla tak samo opisanych przykładów LUX i AUT
- Dla klasy GOOD dolne i górne przybliżenie są równe - klasa ta jest więc zbiorem definiowalnym

Jakość klasyfikacji jest pojedynczym parametrem, który odzwierciedla spójność tabeli decyzyjnej - oblicza się ją jako stosunek sumy licznosci dolnych przybliżeń do licznosci zbioru danych, a więc jakość klasyfikacji wyraża procentowo, co do ilu przykładów nie mamy wątpliwości nt. ich przynależności klasowej (w przykładzie $(2+2+4)/10=8/10$)

Zad. XV. Indukcja i wykorzystanie reguł decyzyjnych

- **Zbiory przybliżone służą do strukturalizacji wiedzy zgromadzonej w tabeli decyzyjnej**
- Następnie poszczególne elementy uzyskane w ramach strukturalizacji (dolne i górne przybliżenia oraz brzegi klas) wykorzystuje się do **indukcji reguł decyzyjnych** - zadaniem reguł jest zwięźle opisanie wiedzy zgromadzonej w tabeli; reguły zastosowane do nowych przykładów (dla których przynależność do klas jest nieznana) mogą pomóc w wypracowaniu dla nich rekomendacji
- Z punktu widzenia decydenta interesujący jest **minimalny zbiór minimalnych reguł decyzyjnych** - minimalny zbiór nie zawiera nadmiarowych reguł (opisujących przykłady już opisane przez inne reguły), minimalna reguła decyzyjna nie zawiera nadmiarowych warunków (bez których i tak pokrywane są te same przykłady)

W kontekście zbioru przybliżonych wykorzystuje się reguły postaci:

jeżeli (koniunkcja warunków elementarnych) to decyzja,

gdzie warunek elementarny ma postać (atrybut = wartość), a decyzja wskazuje na klasę lub klasy

Algorytmy indukcji reguł decyzyjnych działają tak samo niezależnie od tego, na podstawie którego elementu (dolne lub górne przybliżenia albo brzegi klas) są tworzone - różnica polega na interpretacji uzyskanych reguł:

- Reguły **deterministyczne** (pewne) wyznaczamy, dając na wejście algorytmu **dolne przybliżenie** zbioru X - w części decyzyjnej "na pewno klasa X"
- Reguły **możliwe** wyznaczamy, dając na wejście algorytmu **górne przybliżenie** zbioru X - w części decyzyjnej "być może klasa X"
- Reguły **przybliżone** wyznaczamy, dając na wejście algorytmu **brzegi** klas X i Y i ... - w części decyzyjnej "klasa X lub klasa Y lub ..."

Heurystycznym algorytmem wyznaczania minimalnego zbioru reguł jest **LEM2**:

- Przy wyborze warunków do reguły bierze on pod uwagę dwa czynniki w porządku leksykograficznym: liczba pokrywanych przykładów z interesującego nas zbioru (ma być jak największa) oraz liczba pokrywanych przykładów ogółem (ma być jak najmniejsza) (patrz warunki KOSZT=H oraz MOC=L dla klasy BAD - pokrywają po dwa przykłady (BEL,NED) z interesującego zbioru, ale KOSZT=H ogółem pokrywa tylko 3 przykłady ogółem, a MOC=L pokrywa 4 takie przykłady, więc KOSZT=H trafia jako pierwszy do reguły)
- Wybór warunków do reguły przeprowadza się do momentu, w którym pokrywane są tylko przykłady z interesującego nas zbioru (nie zawsze da się pokryć wszystkie przykłady jedną regułą) - np. dla klasy BAD reguła "jeżeli KOSZT=H to BAD" nie jest dobra, bo pokrywa nadmiarowe przykłady; dlatego trzeba do niej dołożyć dodatkowe warunki, rozważając dokładnie te same czynniki wyboru; w tym wypadku warunkiem najlepszym spośród: MOC

= L, BEZP = M oraz BEZP = L okaże się MOC = L, bo ze zbioru wariantów, który pozostał to pokrycia przez regułę, warunek ten pokrywa 2 obiekty, a pozostałe warunki pokrywają tylko 1 obiekt; reguła " jeżeli KOSZT=H oraz MOC = L to BAD" jest już dobra, bo pokrywa tylko obiekty ze zbioru, który chcieliśmy opisać i żadnych nadmiarowych; co więcej, pokrywa wszystkie te obiekty, więc dla klasy BAD nie trzeba szukać dodatkowych reguł;

- Po uzyskaniu pojedynczej reguły należy sprawdzić, czy nie ma ona warunków nadmiarowych (patrz pierwsza reguła deterministyczna dla klasy GOOD, gdzie warunek MOC=M okazał się nadmiarowy; sytuacja taka miała miejsce, gdyż LEM2 jest algorytmem zachłannym);
- Jeśli pojedyncza reguła nie wystarczy do opisu całego zbioru przykładów (dolne lub górnego przybliżenia albo brzegu klas), algorytm stosuje się ponownie tylko do tych, które jeszcze nie zostały pokryte; np. dla klas AVE i GOOD konieczna była indukcja po dwóch reguł; dla klasy AVE - reguła " Jeżeli BEZP = M oraz MOC = H to AVE" pokrywa tylko GER, zaś ITA musi być pokryte inną regułą (w tym wypadku - "jeżeli KOSZT = M oraz BEZP = M to AVE");
- Po uzyskaniu całego zbioru reguł należy sprawdzić, czy nie istnieje jakaś reguła nadmiarowa, tzn. czy przykłady pokrywane przez daną regułę nie są przypadkiem pokrywane przez pozostałe reguły; jeśli by tak była, to taka reguła nie jest nam potrzebna, bo chcemy opisywać wiedzę tak zwięźle jak to tylko możliwe.



Fundusze Europejskie
Polska Cyfrowa



**Rzeczpospolita
Polska**

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego

