

## Dekompozycyjna metoda syntezy lingwistycznych baz wiedzy o badanej zależności.

Opracował: prof. Andrzej Piegat

Synteza baz wiedzy o problemach realizowana jest głównie przez ludzi – zwykle ekspertów problemu. **Lingwistyczna baza wiedzy (LBW)** umożliwia „uczenie komputerów” nie za pomocą wzorów matematycznych i poleceń programowych, lecz z użyciem ludzkiego języka. Człowiek może względnie łatwo utworzyć LBW, jeżeli w problemie występują tylko 2 zmienne niezależne  $x_1, x_2$  (i jedna zmienna zależna  $y$ ) – mamy wówczas do czynienia z problemem 3-wymiarowym (3D).

**Przykładem** takiego problemu jest **problem dozowania środka usypiającego** izofluranu w mieszance oddechowej pacjenta ( $y$ ) zależnego od jego aktualnego ciśnienia krwi ( $x_1$ ) oraz od przeciętnej odchyłki tego ciśnienia względem ciśnienia optymalnego ( $x_2$ ) w dotychczasowym przebiegu operacji.

**Zależności 3-D** człowiek potrafi zrozumieć względnie dobrze. Jeśli jednak wymiar problemu rośnie (**4D, 5D.....**) człowiek gwałtownie traci możliwość zrozumienia zależności występujących w problemie i nie jest w stanie wygenerować lingwistycznej BW.

### Przykłady zależności silnie wymiarowych:

- ✓ zależność globalnej atrakcyjności nieruchomości (dom z ogrodem) od jej atrybutów (cech)  $x_i$ , których liczba może wynosić 18 i więcej, zależnie od tego, które cechy nieruchomości są dla klienta istotne.
- ✓ zależność atrakcyjności ( $y$ ) samochodu od jego cech ( $x_i$ ). Cech, które klient bierze pod uwagę przy zakupie samochodu może być bardzo wiele.
- ✓ zależność atrakcyjności firmy biorącej udział w przetargu na wykonanie potrzebnych nam prac od cech firmy i warunków jakie firma oferuje.
- ✓ zależność atrakcyjności ( $y$ ) sprzętu komputerowego od cech tego sprzętu i potrzeb klienta.

## ZADANIE OCENY NIERUCHOMOŚCI (DOM Z OGRODEM)

Cechy posiadłości wpływające na cenę, jaka gotowi bylibyśmy za nią zapłacić:

C<sub>1</sub> – liczba pokoi

C<sub>2</sub> – powierzchnia użytkowa domu [m<sup>2</sup>]

C<sub>3</sub> – powierzchnia ogrodu [m<sup>2</sup>]

C<sub>4</sub> – liczba garaży

C<sub>5</sub> – wiek domu [lata]

C<sub>6</sub> – liczba łazienek

C<sub>7</sub> – komfort wyposażenia domu {b. mały, mały, średni, duży, b. duży}

C<sub>8</sub> – zewnętrzna, wizualna atrakcyjność domu {BM, M, Sr, D, BD}

C<sub>9</sub> – wizualna atrakcyjność ogrodu {BM, M, Sr, D, BD}

C<sub>10</sub> – stopień zużycia domu i jego instalacji {BM, M, Sr, D, BD}

C<sub>11</sub> – dogodność komunikacyjna usytuowania nieruchomości {BM, M, Sr, D, BD}

C<sub>12</sub> – odległość względem szkół, uczelni, sklepów, ośrodków zdrowia {BM, M, Sr, D, BD}

C<sub>13</sub> – spokój (brak hałasu w otoczeniu), {BM, M, Sr, D, BD}

C<sub>14</sub> – kultura sąsiedztwa {BM, M, Sr, D, BD}

C<sub>15</sub> – bezpieczeństwo dzielnicy {BM, M, Sr, D, BD}

C<sub>16</sub> – bliskość terenów rekreacyjnych (lasów, jezior) {BM, M, Sr, D, BD}

C<sub>17</sub> – wizualna atrakcyjność najbliższego otoczenia {BM, M, Sr, D, BD}

C<sub>18</sub> – techniczne uzbrojenie terenu (woda, kanalizacja, prąd, drogi, etc.) {BM, M, Sr, D, BD}

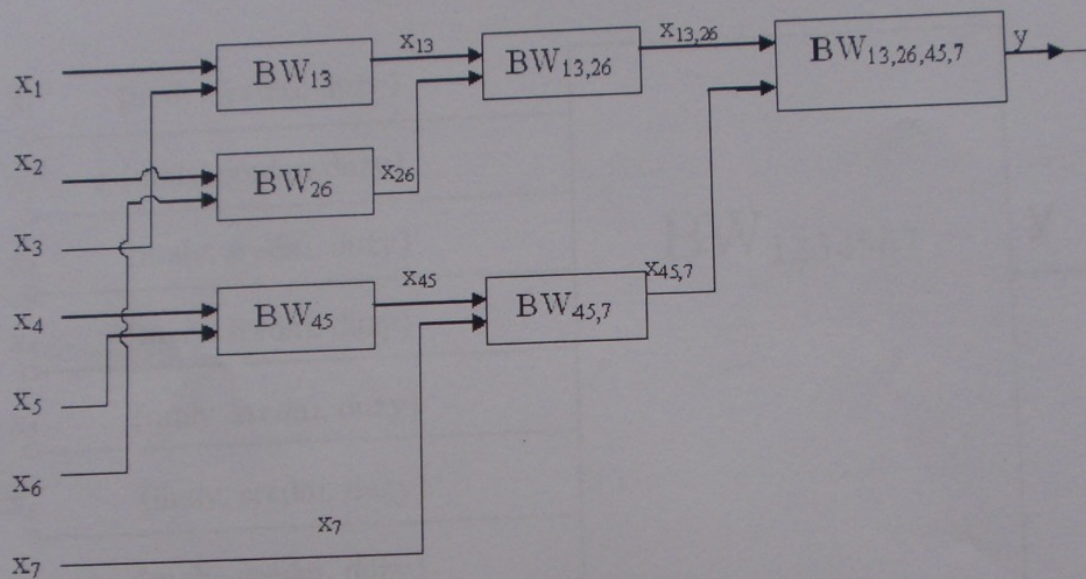
C<sub>19</sub> – występowanie duchów w domu {tak, nie}

**Komentarz:** Większość z cech nieruchomości wpływających na jej wartość to cechy nie liczbowe, lecz cechy mierzone percepcyjnie przez ludzi i wyrażane słowami {BM, M, Sr, D, BD}.

Czy można nie brać tych cech pod uwagę przy ocenianiu wartości nieruchomości? Czy można określić (podać) cenę nieruchomości tylko na podstawie 6 cech liczbowych? Jak wiarygodna byłaby taka ocena?

Jeżeli zmienna  $y$  zależy od wielu zmiennych  $x_i$  to należy dokonać dekompozycji lingwistycznej BW na 3-wymiarowe bazy. W tym celu należy najpierw zagregować zmienne  $x_i$  w pary.

Np.: Jeżeli w problemie występuje 7 zmiennych  $x_i, i=1 \dots$  to agregacja zmiennych może wyglądać tak jak to pokazano na rys.1:



Rys.1. Przykład agregacji 7 zmiennych  $x_i$  i dekompozycji globalnej bazy wiedzy  $BW_{1234567}$  na 6 lokalnych baz wiedzy  $BW_{13}, BW_{26}, BW_{45}, BW_{13,26}, BW_{45,7}$  i  $BW_{13,26,45,7}$ .

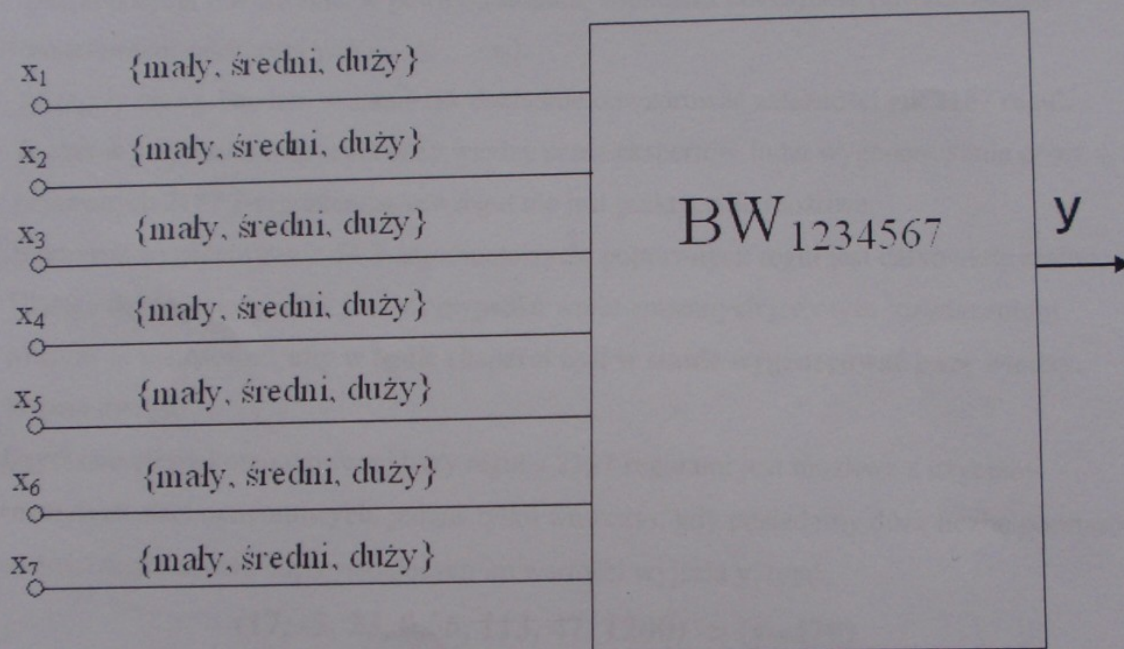
### Jakie korzyści daje agregacja zmiennych (dekompozycja) globalnej bazy wiedzy?

1. Pierwszą korzyścią jest ułatwienie ekspertowi generowania bazy wiedzy. Ekspert generuje reguły z przestankami 2-argumentowymi (łatwymi do zrozumienia) a nie z przestankami n-argumentowymi (trudnymi do zrozumienia).
2. Dzięki dekompozycji globalnej bazy reguł w lokalne bazy znacznie spada całkowita liczba reguł.

**Przykład:**

Jeżeli liczba  $n$  zmiennych niezależnych  $x_i$  wynosi 7, a każda ze zmiennych oceniana będzie z użyciem 3 kwantyfikatorów lingwistycznych „mały, średni, duży” to w niezdekomponowanej bazie wiedzy wystąpi:

$$3*3*3*3*3*3*3 = 3^7 = 2187 \text{ reguł.}$$



Przykładowe reguły:

**IF** ( $x_1 = \text{duży}$ ) **AND** ( $x_2 = \text{mały}$ ) **AND** ( $x_3 = \text{średni}$ ) **AND** ( $x_4 = \text{mały}$ ) **AND** ( $x_5 = \text{duży}$ )  
**AND** ( $x_6 = \text{średni}$ ) **AND** ( $x_7 = \text{mały}$ ) **THEN** ( $y = ?$ )

Rys. 2. Niezdekomponowana baza wiedzy o zależności  $y = f\{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$  i przykładowa reguła z tej bazy, w której należy zdefiniować konkluzje ( $y=?$ )

W przypadku zdekomponowania globalnej bazy wiedzy w bazy lokalne całkowita liczba reguł wynosić będzie, przy założeniu, że wszystkie zmienne wejściowe  $x_i$  i wewnętrzne oceniane będą z użyciem 3 kwantyfikatorów {mały, średni, duży}:

$$3 * 3 + 3 * 3 + 3 * 3 + 3 * 3 + 3 * 3 + 3 * 3 =$$

$$9 + 9 + 9 + 9 + 9 + 9 = 54$$

Liczba reguł zmniejsza się, więc z ogromnej liczby **2187 trudnych** do zrozumienia 7-argumentowych reguł do **54 łatwych** do zrozumienia reguł 2-argumentowych typu:

**IF (x<sub>1</sub> duze) AND (x<sub>2</sub> srednie) THEN (y=?).**

### **Czy dekompozycja lingwistycznych baz wiedzy ma same zalety?**

Dekompozycja BW zwykle, w pewnym stopniu, zmniejsza dokładność odwzorowania modelowanej zależności  $y=f(x_1, x_2, \dots, x_n)$ .

54 reguły nie są, bowiem w stanie tak dokładnie odwzorować zależności jak 2187 reguł.

Jednak w przypadku tworzenia bazy wiedzy przez ekspertów-ludzi wygenerowanie przez nich poprawnych 2187 7-argumentowych reguł nie jest praktycznie możliwe.

Natomiast wygenerowanie 54 2-argumentowych, poprawnych reguł jest całkowicie realne.

Dlatego dekompozycja BW jest w przypadku wielu zmiennych jedynym rozwiązaniem

Musimy ją zastosować, **aby w ogóle eksperci byli w stanie wygenerować bazę wiedzy.**

#### **Ważna uwaga:**

Uzyskanie niezdekomponowanej bazy reguł z 2187 regułami jest możliwe z użyciem **rozmytych sieci neuronowych**, jednak tylko wówczas, gdy posiadamy dużą liczbę pomiarów wartości wejść  $x_i$  oraz odpowiadających im wartości wyjścia  $y$ , typu:

$$(17, -5, 23, 0,56, 113, 47, 1200) \rightarrow (y=479)$$

Takich pomiarów badanej zależności potrzeba kilka razy więcej niż wynosi liczba reguł (2187). Jednak w praktyce pomiarów zwykle brak lub ich liczba jest za niska, aby umożliwić wygenerowanie BW.

#### **Zasady agregacji zmiennych niezależnych $x_i$ :**

Wśród zmiennych  $x_i$  wpływających na zmienna zależną  $y$  są zwykle zmienne zwane **stymulantami** oraz **destymulanty**. Przykładem stymulanty jest powierzchnia domu [ $m^2$ ]. Z jej wzrostem rośnie atrakcyjność domu dla klienta (zmiennej zależnej  $y$ ). Natomiast wrak domu (samochód) jest destymulantą. Z jego wzrostem atrakcyjność domu (samochodu) maleje.

W trakcie agregacji korzystnie jest łączyć ze sobą w pary zmienne  $x_i$  tego samego typu: stymulanty ze stymulantami, destymulanty z destymulantami. Nie musi to być jednak

bezwzględna zasada. Jeżeli połączenie pewnej stymulanty z destymulantą umożliwia łatwe tworzenie reguł dzięki dobremu zrozumieniu ich wpływu można odstąpić od powyższej zasady.

Jeżeli zastosujemy metodę oddzielnej agregacji stymulant i destymulant to ostatecznie dojdziemy do końcowej, lokalnej bazy wiedzy w formie:

$$y=f(\text{zagregowane stymulanty, zagregowane destymulanty}).$$

Reguły lingwistyczne definiujące powyższą zależność będą łatwe do określenia przez ekspertów.

### **Przykład generowania zdekomponowanej bazy wiedzy w problemie oceny i wyboru najlepszej firmy biorącej udział w przetargu na wykonanie prac.**

Ogłaszamy przetarg na wykonanie pewnych prac (np.: informatyzacja firmy, budowa autostrady, remont parku maszynowego, etc.)

Do przetargu zgłasza się szereg firm oferujących wykonanie prac za pewną cenę  $x_{li}$  [tys. zł]

Jedne firmy gotowe są wykonać prace za niższą inną za wyższą cenę. Jednak wybór najtańszej firmy może okazać się dużym błędem, ponieważ firma ta może wykonać zleconą jej pracę źle. Obowiązuje znana prawda: zwykle najtańszy towar nie jest najlepszy. Czasem nie nadaje się on w ogóle do użytku. Dlatego przy wyborze firmy należy ocenić jej całkowitą atrakcyjność  $y$  jako oferenta. Na atrakcyjność oferenta składa się wiele cech, przy czym, niektóre cechy mają wyższą istotność inne niższą. Istotność danej cechy zależy od poglądów (**preferencji**) osoby oceniającej oraz od czynników obiektywnych.

Np.: dla jednej osoby **wygląd samochodu** może mieć większą istotność niż jego marka. Dla innej może być najważniejsza.

W przypadku **budowy stadionu na mistrzostwa świata** w piłce nożnej cena może mieć mniejsze znaczenie niż termin zakończenia budowy stadionu, który musi być bezwzględnie dotrzymany. Terminowość wykonania prac przez oferenta będzie, więc najważniejsza. Aby stworzyć lingwistyczna BW do oceny atrakcyjności  $y$  poszczególnych oferentów należy najpierw określić wszystkie atrybuty (cechy)  $x_i$ , które wpływają na ogólną atrakcyjność oferenta.

**UWAGA:** cechami  $x_i$  mogą być tylko takie **cechy, których wartości są osiągalne**. Jeżeli wartości jakiejś cechy  $x_i$ , potencjalnie istotnej, nie sposób znikąd zdobyć (brak informacji o jej wartości), nie można jej użyć w systemie oceny.

Wprowadzimy teraz następujące pojęcia:

1. Bezwzględna atrakcyjność oferenta  $A_B$
2. Względna atrakcyjność oferenta  $A_w$ .

Atrakcyjność bezwzględna będzie bezpośrednią funkcją cech oferenta:

$$A_B = f(x_1, x_2, \dots, x_n)$$

W przybliżeniu atrakcyjność bezwzględna jest proporcjonalna do ważonej różnicy zalet i wad danego oferenta.  $A_B \approx (\text{ważone zalety} - \text{ważone wady})$ .

Atrakcyjność względna oferenta będzie ilorazem atrakcyjności bezwzględnej i ceny, jaką oferent Ada za swoje usługi.

$$A_w = \frac{A_B}{\text{cena [tys. zł]}}$$

Atrakcyjność względna będzie nam, więc informowała ile atrakcyjności bezwzględnej uzyskujemy za każde 1000zł, które ewentualnie zapłacimy danemu oferentowi.

W uproszczeniu i przybliżeniu można powiedzieć, że atrakcyjność względna jest proporcjonalna do różnicy wad i zalet danego oferenta przypadającej na każdy 1000zł ceny.

$$A_w \approx \frac{(\text{ważone zalety} - \text{ważone wady})}{\text{cena [tyś. zł]}}$$

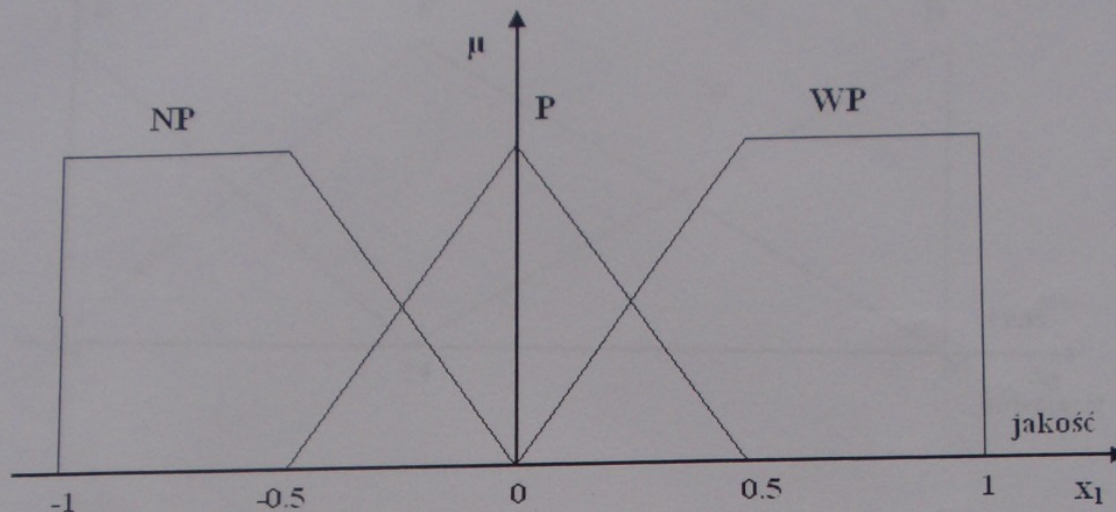
Atrakcyjność oferenta jest, więc „towarem”, który kupujemy. Za każde wydane 1000 zł. chcemy uzyskać jak najwięcej tego „towaru”.

Po namyśle doszliśmy do wniosku, że o atrakcyjności oferenta świadczą jego następujące cechy:

$x_1$  – **jakość realizacji prac**. Jakość tą możemy ocenić badając, jak dotychczas dany oferent wykonywał prace zlecane mu przez inne firmy. Jakość  $x_1$  oceniać będziemy w zakresie  $[-1, 0, +1]$ , gdzie  $x_1 = 1$  oznacza najwyższą jakość powyżej przeciętnej, „-1” – najgorszą, złą jakość (poniżej przeciętnej), „0” oznacza jakość przeciętną.

Jakość  $x_1$  realizacji prac oceniana będzie z użyciem następujących ocen lingwistycznych: niżej przeciętnej NF, przeciętna P, wyżej przeciętnej WP.

Funkcje cechy tych kwantyfikatorów podano na rys. 3.



Rys.3. Funkcje cechy jakości realizacji prac przez oferenta.

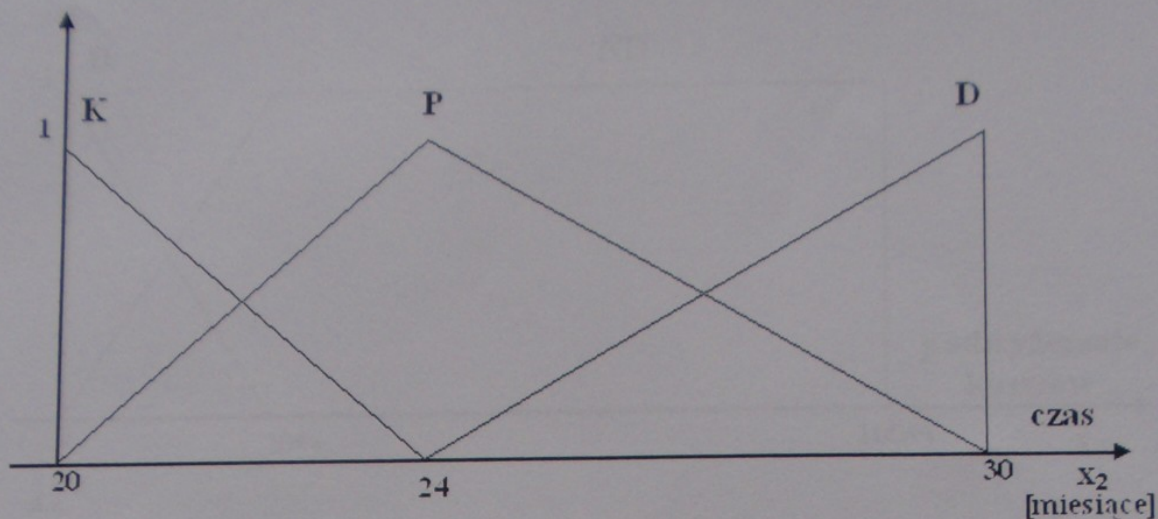
Funkcje z rys. 3 wyrażają pogląd, że jeżeli jakość prac realizowanych przez danego oferenta przekracza o 50% przeciętną jakość, dany oferent uznawany jest w 100% za „wyżej niż przeciętnego”. Podobny sens dotyczy wartości  $x_1 \leq -0.5$

**$x_2$  – czas wykonywania prac(oferty) [miesiące] deklarowany przez oferenta.**

Każdy ze zgłaszających się oferentów deklaruje pewien czas realizacji prac. Można więc obliczyć przeciętnie oferowany czas wykonania oferty i z czasem tym porównywać wszystkie oferty.

Założmy, że w przypadku otrzymanych przez nas ofert proponowane czasy oscylują w zakresie 20 do 30 miesięcy, a **przeciętny czas (średnia ze wszystkich ofert) wynosi 24 miesiące**. Na podstawie tych informacji określamy funkcje cech kwantyfikatorów **Krótki czas (K), przeciętny czas (P), długi czas (D)**, pokazane na rys.4.



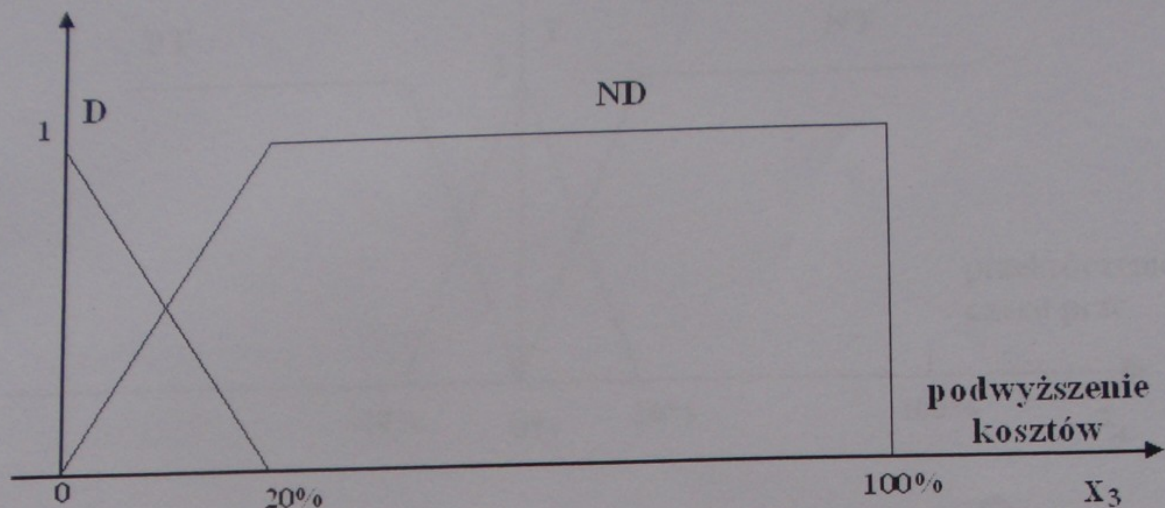


Rys.4. Funkcje cechy kwantyfikatorów lingwistycznych

krótki czas (K), przeciętny czas (P) i długi czas (D) realizacji prac przez oferenta.

### **$x_3$ – dotrzymywanie (podwyższanie) kosztów prac przez oferenta.**

Często zdarza się, że oferent poda w ofercie zbyt niską cenę wykonania prac, po to, aby wygrać przetarg. Jednak potem, gdy nie potrafi zrealizować wszystkich prac po deklarowanych kosztach żąda „dopłacenia” pewnej sumy do pierwotnie uzgodnionej ceny oferty. Czasem jednak podwyższenie kosztów prac wynika nie z błędnego ich oszacowania, lecz z nieprzewidzianych sytuacji i wypadków, jakie mogą wystąpić w trakcie realizacji oferty (np. spalenie się świeżo zbudowanego budynku). Dlatego zleceniodawca powinien zaakceptować ewentualne podwyższenie kosztów względem ustalonych kosztów, o ile podwyższenie to nie jest zbyt duże (wg opinii zleceniodawcy). Na rys.5 przedstawiono funkcje cechy kwantyfikatorów lingwistycznych: oferent „dotrzymuje koszty (D)”, „nie dotrzymuje kosztów (ND)”.

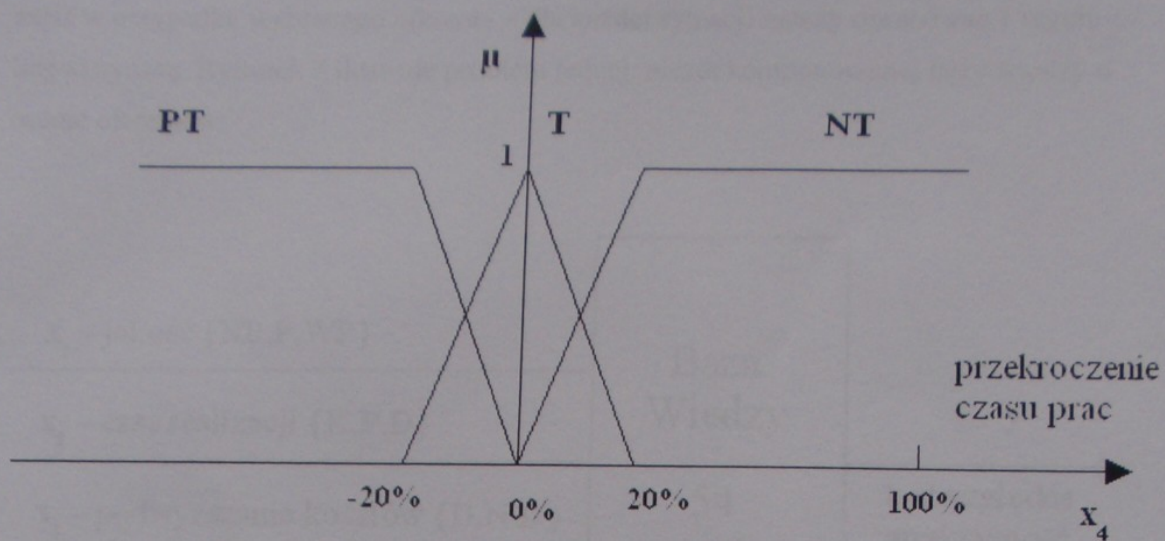


Rys.5. Funkcje cechy kwantyfikatorów D - dotrzymuje kosztów, oraz ND - nie dotrzymuje kosztów, służących do oceny zmiennej  $x_3$  - niedotrzymywanie (podwyższanie) kosztów przez oferenta w trakcie realizacji prac.

#### $x_4$ – terminowość realizacji prac przez oferenta.

Oferent może być znany z tego, że często przedłuża czas realizacji prac powyżej deklarowanego i uzgodnionego w umowie (**NT-nieterminowy**), wykonuje prace zwykle w terminie (**T-terminowy**), lub często wykonuje prace przed terminem (**PT-przedterminowy**). Za terminowego oferenta uważa się również takiego oferenta, który choć jest czasem nieterminowy, to odchyłki terminu są małe.

Na rys.6 przedstawione są funkcje cechy kwantyfikatorów oceny cechy  $x_4$ .



Rys. 6. Funkcje cechy kwantyfikatorów do lingwistycznej oceny przekroczenia czasu realizacji zlecenia.

Założmy, że więcej cech oferentów nie będziemy uwzględniali, chociaż można by dodać kolejne, np.  $x_5$  – solidność przyjmowania i wykonywania reklamacji przez oferenta.

**Ustalenie rankingu cech  $x_i$  wg. stopnia ich istotności dla nas, jako dla zleceniodawcy.**

Po zastanowieniu się ustaliliśmy następujący ranking cech  $x_i$  (od najważniejszej do najmniej ważnej cechy):

$x_1$  – jakość realizacji prac (cecha najważniejsza)

$x_2$  – czas wykonania prac

$x_3$  – niedotrzymywanie (podwyższanie) kosztów zlecenia

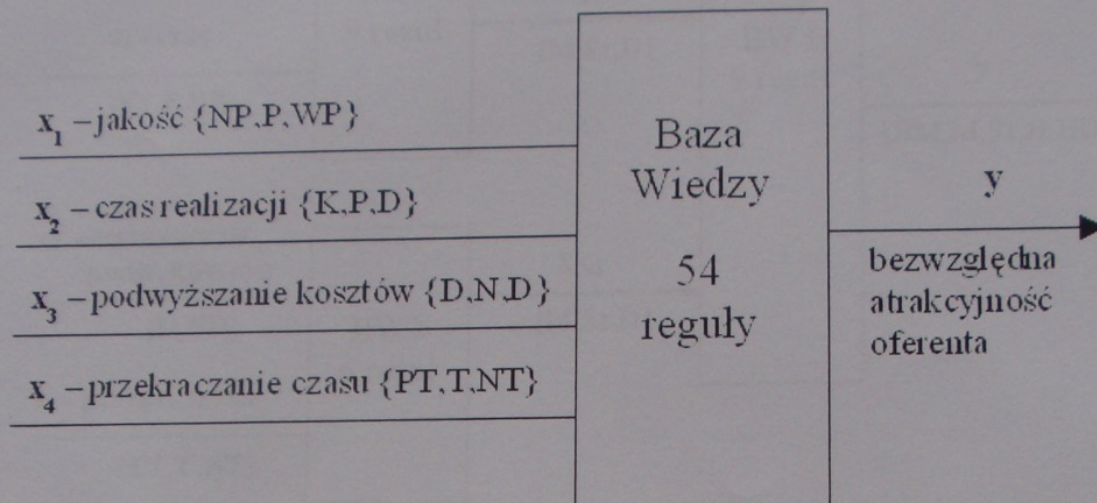
$x_4$  – terminowość realizacji zlecenia (cecha najmniej ważna w sensie względnym – nie znaczy to jednak, że w ogóle nieważna).

**Uwaga**

Atrybut  $x_1$  oferenta oceniany jest przez 3 kwantyfikatory,  $x_2$  - przez 3,  $x_3$  - przez 2, a  $x_4$  przez 3 kwantyfikatory. W związku z tym liczba możliwych kombinacji wartości tych kwantyfikatorów wynosi

$$3 \times 3 \times 2 \times 3 = 54$$

Każda z możliwych kombinacji wartości lingwistycznych opisuje jedną sytuację, jaka może zajść w przypadku wybranego oferenta – dla każdej sytuacji należy opracować 1 regułę lingwistyczną. Rysunek 7 ilustruje problem jednej, niezdekomponowanej bazy wiedzy o ocenie oferentów.



Rys.7. Niezdekomponowana baza wiedzy do oceny bezwzględnej atrakcyjności oferenta z 54 regułami.

W przypadku niezdekomponowanej BW przesłanki wszystkich reguł będą 4-składnikowe. Poniżej przykłady 2 reguł:

**IF** (jakość przeciętna) **AND** (czas realizacji przeciętny) **AND** (koszty dotrzymuje)  
**AND** (czas realizacji nieterminowy) **THEN** (ocena = ?)

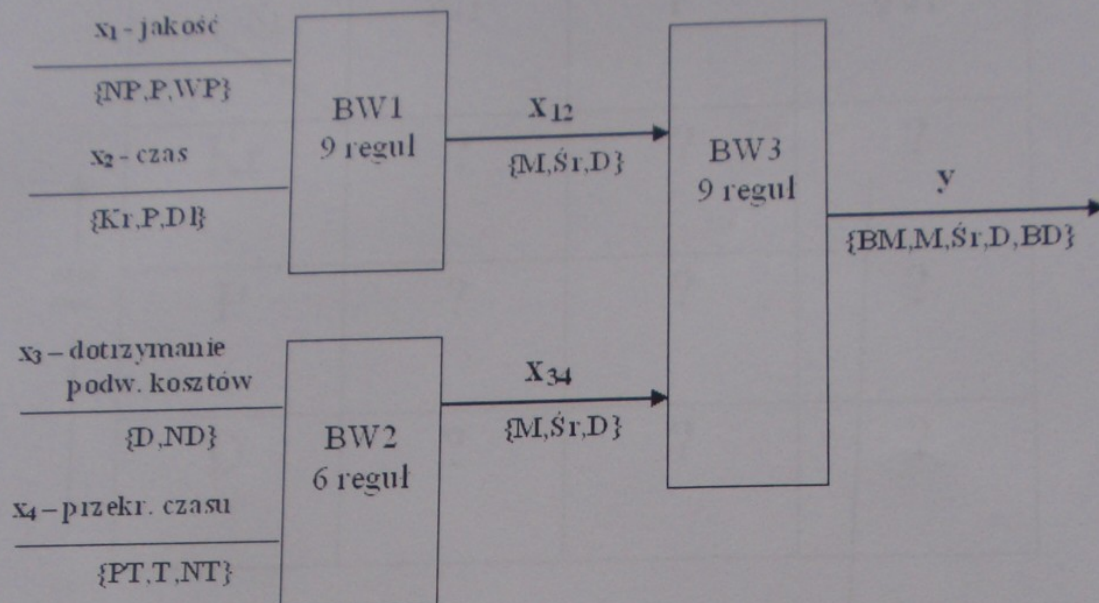
**IF**(jakość wyżej przeciętnej) **AND** (czas realizacji długi) **AND** (koszty dotrzymuje)  
**AND** (czas realizacji nieterminowy) **THEN** (ocena = ?)

Stosując do oceny atrakcyjności oferenta 5 kwantyfikatorów:

{bardzo wysoka BW, wysoka W, średnia Śr, niska N, bardzo niska BN},

trudno jest zdecydować jakie powinny być **konkluzje** w powyższych 2 regułach, **czy mają być takie same czy zróżnicowane**, a jeśli tak to jakie?

W dalszym ciągu dokonana zostanie dekompozycja globalnej bazy wiedzy na 3 bazy składowe. Plan dekompozycji przedstawiony jest na rys.8



Rys.8. Schemat dekompozycji globalnej bazy wiedzy na 3 bazy składowe – suma reguł 24.

Taki plan dekompozycji pozwala **zmniejszyć liczbę reguł** ponad 2-krotnie:

**z 54 do 24**

### 1-sza, częściowa baza wiedzy (BW1)

W bazie tej zostaną zagregowane dwa najistotniejsze atrybuty oferenta:

**x<sub>1</sub> – jakość realizacji prac** {niżej przec., przec., wyżej przec.}

**x<sub>2</sub> – czas realizacji prac** {krótki, przeciętny, długi}

Wielkością wynikową **x<sub>12</sub>** będzie **łączna ocena** jakości i czasu realizacji prac **x<sub>12</sub>** (jakość, czas). Wartość tej łącznej oceny wyrażona będzie z użyciem 3 kwantyfikatorów lingwistycznych

**x<sub>12</sub>: {mała, średnia, duża}**

Na rys. 9 pokazana jest pusta, niewypełniona jeszcze konkluzjami BW1 w formie tabelarycznej.

jakość prac →

$x_2$	$x_1$	NP	P	WP
Kr	?	?	?	?
P	?	?	?	?
D	?	?	?	?

↑  
czas  
prac

wartość lingwistycznej  
zagregowanej cechy  
 $x_{12}$  (jakość, czas)

Rys.9. Pusta baza wiedzy BW1 z 9 możliwymi kombinacjami cech  $x_1$  (jakość) i  $x_2$  (czas prac), dla których należy określić konkluzje: lingwistyczne wartości zagregowanej oceny  $x_{12}$  (jakość, czas).

Do opracowania BW1 zastosowana zostanie **metoda rankingowa**. Cecha  $x_1$  (jakość prac) jest **stymulantą** łącznej oceny, cecha  $x_2$  (czas prac) **destymulantą**. W związku z tym najlepsze, 1-sze miejsce w rankingu musi zająć kombinacja:

**1:** [ $x_1$  (jakość prac)=wyżej przeciętnej]AND[ $x_2$  (czas prac)=krótki]

natomiast najniższe 9-te miejsce w rankingu musi zająć kombinacja:

**9:** [ $x_1$  (jakość prac)=niżej przeciętnej]AND[ $x_2$  (czas prac)=długi]

Następnie należy rozważyć i porównać ze sobą wszystkie pozostałe kombinacje i ustalić ich ranking, uwzględniając istotność poszczególnych cech:

$x_1$  (jakość) – istotność wyższa

$x_2$  (czas) – istotność niższa

Np. należy rozważyć, która z poniższych kombinacji jest dla nas (zleceniodawcy) korzystniejsza:

$[x_1$  (jakość) = przeciętna] AND (czas=krótki)

$[x_1$  (jakość) = wyżej przeciętnej] AND (czas=średni)

Na rys. 10 przedstawiono kolejność rankingową (1 do 9) poszczególnych kombinacji wartości cech  $x_1, x_2$  oferenta.

jakość prac →

$x_1 \backslash x_2$	NP	P	WP
Kr	6	3	1
P	8	5	2
D	9	7	4

↑ czas prac

pozycja w ranking

Rys.10. Tabela rankingowa kombinacji wartości cech  $x_1$  (jakość prac) oraz  $x_2$  (czas realizacji prac), 1-pozycja najwyższa, 9-najniższa.

Następnie należy przyporządkować poszczególnym kombinacjom wartości lingwistycznych  $x_1$ ,  $x_2$  lingwistyczne wartości łącznych ocen  $x_{12}$  mając do wyboru 3 oceny:

$x_{12} \dots \{\text{mała, średnia, duża}\}$

Przyporządkowanie podane jest poniżej:

- 1 – duża, D
- 2 – duża, D
- 3 – średnio-duża, Śr-D
- 4 – średnio-duża, Śr-D
- 5 – średnia, Śr
- 6 – mało-średnia, M-Śr
- 7 – mało-średnia, M-Śr
- 8 – mała, M
- 9 – mała, M

Określenie „średnio-duża” oznacza, że łączna ocena  $x_{12}$  jest pośrednia między oceną średnią i dużą.

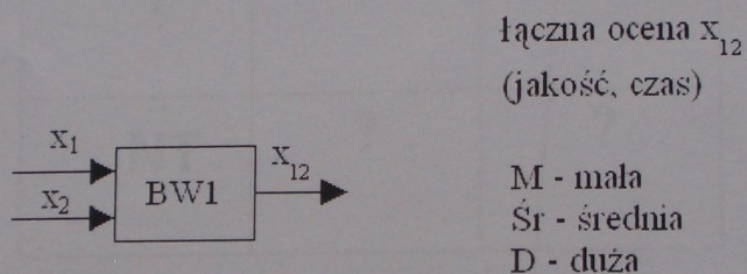
Na rys. 11 przedstawiono gotową BW1 z podanymi lingwistycznymi zagregowanymi ocenami  $x_{12}$  (jakość, czas)



jakość prac →

$x_2$	$x_1$	NP	P	WP
Kr		M-Śr	Śr-D	D
P		M	Śr	D
D		M	M-Śr	Śr-D

↑  
czas  
prac



Rys. 11. Składowa baza wiedzy BW1 agregująca oceny  $x_1$  (jakość) oraz  $x_2$  (czas realizacji) w jedną ocenę łączną  $x_{12}$  (jakość, czas).

Przykładowa (jedna z 9) reguł wynikająca z BW1:

**IF** [ $x_1$  (jakość)=WP] **AND** [ $x_2$  (czas)=Kr] **THEN** [ $x_{12}$  (jakość, czas)=D]

## 2-ga częściowa baza wiedzy BW2

W bazie BW2 zagregowane zostaną 2 cechy oferenta

$x_3$  – **dotrzymanie kosztów prac** (dotrzymuje D, nie dotrzymuje ND)

$x_4$  – **przekraczanie czasu prac** (przed-terminowy PT, terminowy T, nieterminowy NT)

Baza BW2 powinna być określona z uwzględnieniem faktu, że dla zleceniodawcy  $x_3$  (dotrzymanie kosztów) ma większą istotność niż  $x_4$  (przekraczanie uwzględnionego czasu realizacji).

Na rys. 12 pokazana jest pusta baza wiedzy BW2.

		← dotrzymanie kosztów	
		$x_3$	
$x_4$	D	ND	
↑ przechr. czasu prac	PT	?	?
	T	?	?
	NT	?	?
		łączna ocena $x_{34}$ (dotrz. kosztów, przechr. czasu)	

Rys. 12. Pusta baza wiedzy BW2 agregująca  $x_3$  (dotrzymanie kosztów) i  $x_4$  (przekraczanie czasu prac) obejmująca 6 kombinacji różnych wartości lingwistycznych zmiennych  $x_3$  i  $x_4$ .

Po analizie porównawczej ustalono ranking poszczególnych kombinacji lingwistycznych ocen cechy  $x_3$  oraz  $x_4$  oferenta przedstawiony na rys. 13.

		← dotrzymany e kosztów	
	$x_3$	<b>D</b>	<b>ND</b>
$x_4$	↙		
	<b>PT</b>	<b>1</b>	<b>3</b>
↑ dotrym terminów	<b>T</b>	<b>2</b>	<b>5</b>
	<b>NT</b>	<b>4</b>	<b>6</b>

1 – najwyższa pozycja  
6 – najniższa pozycja  
w rankingu

Rys. 13. Baza BW2 z numerami rankingowymi przyporządkowanymi poszczególnym kombinacjom lingwistycznych wartości cech  $x_3$  (dotrzymany e kosztów) oraz  $x_4$  (dotrzymany e terminów) przez oferenta.

W celu ustalenia rankingu należało zastanowić się, czy np. korzystniejszy jest oferent o cechach:

$[x_3 \text{ (dotrzymany e kosztów)=dotrzymuje}] \text{AND} [x_4 \text{ (dotrzymany e terminów)=terminowy}]$

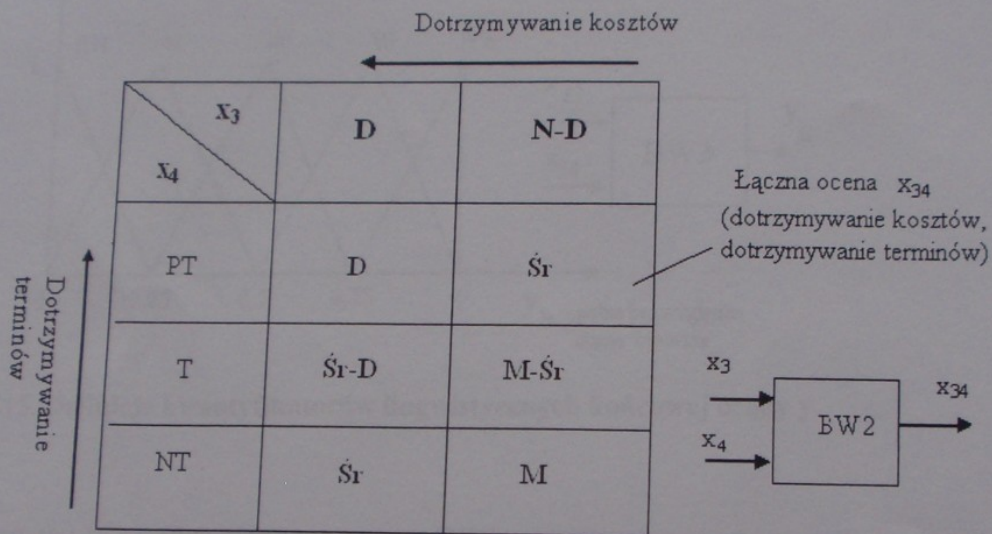
czy

$[x_3 \text{ (dotrzymany e kosztów)=nie dotrzymuje}] \text{AND} [x_4 \text{ (dotrzymany e terminów)=przed-terminowy}]?$

Po ustaleniu rankingu poszczególnych kombinacji cech przyporządkowane zostały im lingwistyczne wartości {mała, średnia, duża} łącznej oceny  $x_{34}$  (dotrzymany e kosztów, dotrzymany e terminów)

- 1 – duża, D
- 2 – średnio-duża, Śr-D
- 3 – średnia, Śr
- 4 – średnia, Śr
- 5 – mało-średnia, M-Śr
- 6 – mała, M

Rys.14 przedstawia wypełnioną konkluzjami gotową bazę wiedzy BW2.



Rys.14 Baza wiedzy BW2 agregująca cechy oferenta  $x_3$  (dotrzymanie kosztów) i  $x_4$  (dotrzymanie terminów).

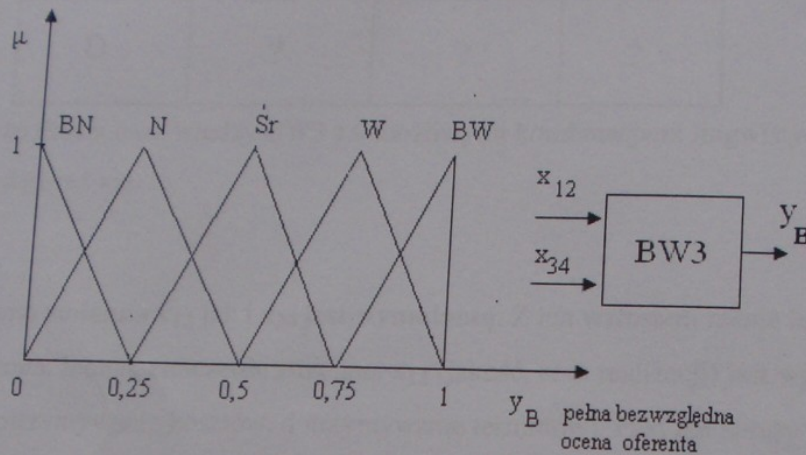
Uwaga !

Ocena zagregowana  $x_{34}$  jest stymulantą. Jeżeli jej wartość lingwistyczna rośnie, rośnie też globalna ocena  $y = x_{1234}$  oferenta.

### Trzecia, częściowa baza wiedzy BW3

W bazie BW3 nastąpi agregacja łącznych cech  $x_{12}$  (jakość, czas realizacji) oraz  $x_{34}$  (dotrzymywanie kosztów, dotrzymywanie terminów) pochodzących z częściowych baz BW1 i BW2, w ostateczną łączną ocenę  $y = x_{1234}$  oferenta.

Na rys.15 przedstawiono definicje 5 lingwistycznych kwantyfikatorów, które będą stosowane względem zmiennej  $y_B$ .



Rys.15. Definicje kwantyfikatorów lingwistycznych końcowej oceny  $y$ .

Na rysunku 16 przedstawiono prostą bazę BW3 zawierającą wszystkie możliwe kombinacje lingwistycznych kombinacji lingwistycznych wartości zmiennej  $x_{12}$  oraz  $x_{34}$ .

$x_{12}$	M	Śr	D	
$x_{34}$	M	Śr	D	
	?	?	?	
	?	?	?	
	?	?	?	

$x_{12}$  - większa istotność  
 $x_{34}$  - mniejsza istotność  
 Wartość łącznej końcowej oceny  $y_B$  oferenta

Rys.16 Prosta baza wiedzy BW3 z 9 możliwymi kombinacjami lingwistycznych ocen  $x_{12}$  oraz  $x_{34}$ .

Zarówno zmienna  $x_{12}$  jak i  $x_{34}$  jest stymulantą. Z ich wzrostem rośnie łączna ocena  $y$  oferenta. Jednak **znaczenie** zmiennej  $x_{12}$  (jakość, czas realizacji) jest **wyższe** niż  $x_{34}$  (dotrzymanie kosztów, dotrzymanie terminów). Fakt ten należy uwzględnić przy ustaleniu rankingu poszczególnych kombinacji. Na pewno **najlepszym oferentem będzie** oferent o cechach:

**( $x_{12} = \text{duże}$ ) AND ( $x_{34} = \text{duże}$ ) - miejsce 1 w rankingu**

Najgorszym oferentem jest oferent o cechach :

**( $x_{12} = \text{małe}$ ) AND ( $x_{34} = \text{małe}$ ) - miejsce 9 w rankingu**

Tabela na rys. 17 przedstawia ranking poszczególnych kombinacji cech  $x_{12}$  i  $x_{34}$  oferenta opracowany z uwzględnieniem różnej istotności  $x_{12}$  i  $x_{34}$ .

		jakość, czas realizacji		cecha ważniejsza	
		→			
Dotrzymanie kosztów terminów		$x_{12}$	M	Śr	D
		$x_{34}$	M	Śr	D
↓		M	9	6	4
		Śr	8	5	2
		D	7	3	1
		jakość, koszty			

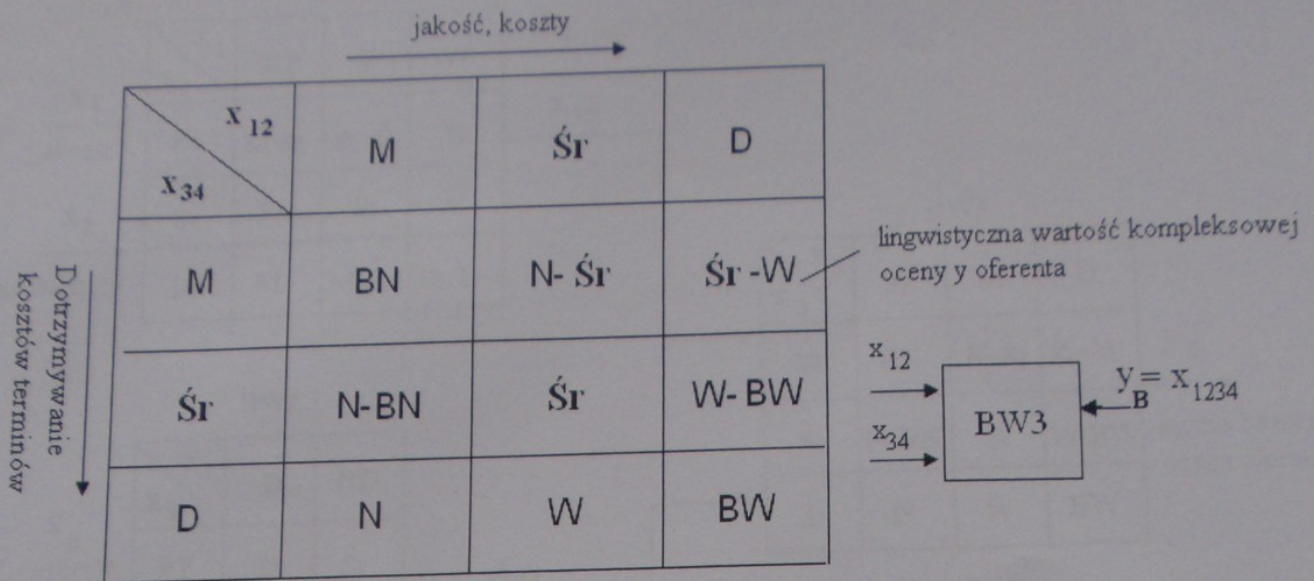
pozycja rankingowa danej kombinacji  $x_{12}, x_{34}$

Rys. 17 Baza wiedzy BW3 z naniesionymi rankingowymi pozycjami poszczególnych kombinacji cech.

Poszczególnych pozycjom rankingowym podporządkowane zostają lingwistyczne kwantyfikatory łącznej oceny  $y_B$  jak poniżej

Ranking	Lingwistyczny kwantyfikator oceny $y$	
1	Bardzo wysoka	BW
2	Wysoka-bardzo wysoka	W-BW
3	Wysoka	W
4	Średnio-wysoka	Śr-W
5	Średnia	Śr
6	Nisko-średnia	N-Śr
7	Niska	N
8	Niska-bardzo niska	N-BN
9	Bardzo niska	BN

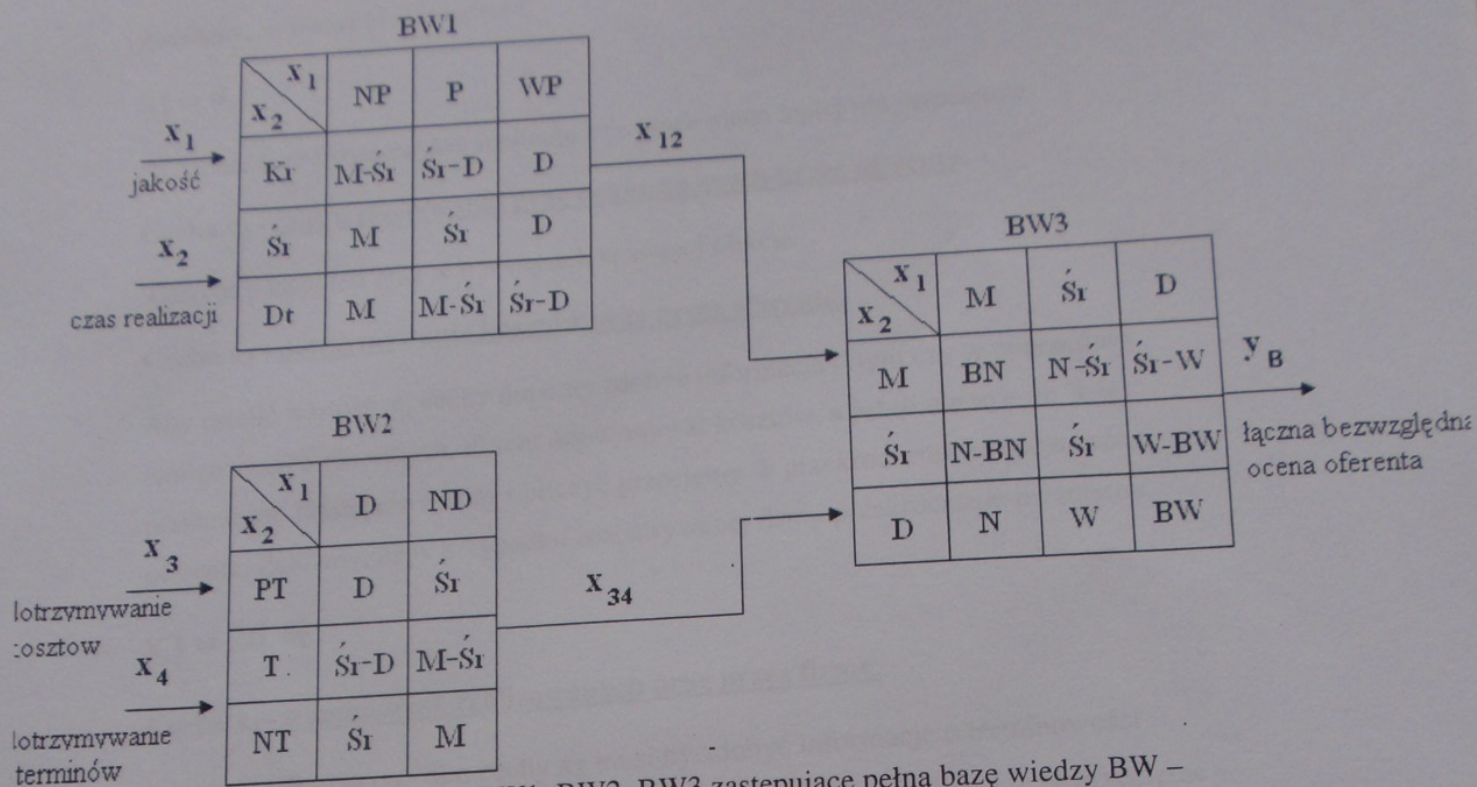
Na rys. 18 przedstawiona jest baza BW3.



**Rys.19** Baza Wiedzy BW3 z regułami przyporządkowującymi kombinacjom lingwistycznych wartości zmiennych  $x_{12}$  oraz  $x_{34}$  –lingwistyczne wartości łącznej oceny y oferenta.

Na rys.20 przedstawiono schemat połączeń wszystkich 3 baz.





Rys.20 Trzy bazy wiedzy BW1, BW2, BW3 zastępujące pełną bazę wiedzy BW – problem bezwzględnej oceny  $y_B$  oferenta w przetargu.

## Jak w praktyce przeprowadzać obliczenia liczbowej oceny oferentów?

Każdy z oferentów musi być oceniony oddzielnie. Rozpatrując pojedynczego, np. pierwszego oferenta musimy **ocenić liczbowo** wartość każdej z 4-ch jego cech.

**Cecha  $x_1$  - jakość prac** realizowanych przez oferenta może zmieniać się zgodnie z przyjętą skalą w zakresie  $[-1,0,1]$ . Zero oznacza przeciętną jakość prac w porównaniu z innymi firmami, +1 oznacza wybitnie wyróżniającą się jakość prac, a -1 oznacza wybitnie złą jakość prac.

Aby ocenić liczbowo jakość prac danej firmy musimy zebrać jak najwięcej informacji o pracach dotychczas realizowanych przez tą firmę i następnie wystawić ocenę liczbową.

Załóżmy, że ocena ta ma wartość:

$$x_1 = 0,1$$

co oznacza, że firma zwykle realizuje swe prace nieco lepiej niż przeciętnie.

**Cecha  $x_2$  - czas wykonywania prac deklарowanych przez oferenta.**

Oferent podaje ten czas w miesiącach w swojej ofercie.

**Cecha  $x_3$  - dotrzymywanie kosztów prac przez oferenta.**

Aby ustalić wartość tej cechy musimy zdobyć informacje o tym czy w poprzednio realizowanych zleceniach oferent dotrzymywał kosztów, a jeżeli nie to o ile % je przekraczał. Następnie należy obliczyć przeciętny % przekroczenia kosztów przez oferenta. Załóżmy, że w przypadku rozpatrywanej firmy przekroczenie przeciętne:

$$x_3 = 20 \%$$

**Cecha  $x_4$  - terminowość realizowanych prac przez firmę.**

Podobnie jak w przypadku cechy  $x_3$  musimy zdobyć informacje o terminowości realizacji poprzednich prac wykonywanych przez oferenta i obliczyć w % przeciętne wydłużanie lub skracanie terminów.

Załóżmy, że rozpatrywany oferent wydłuża przeciętnie czas realizacji prac o 25%:

$$x_4 = 25 \%$$

Posiadając liczbowe wartości cech oferenta możemy je wprowadzić do lingwistycznej bazy wiedzy, która po przeprowadzeniu operacji fuzyfikacji, inferencji i defuzyfikacji obliczy nam dla rozpatrywanego oferenta łączną, bezwzględną ocenę  $y_B$  oferenta.

Załóżmy że obliczona **wartość oceny bezwzględnej oferenta I** wynosi :

$$y_{B1} = 0.37$$

natomiast **oferenta II** wynosi :

$$y_{B2} = 0.43$$

Oferent I żąda za wykonanie prac sumy:

$$C_I = 1.250 \text{ mln zł}$$

Oferent II żąda sumy:

$$C_{II} = 1.350 \text{ mln zł}$$

Aby porównać obu oferentów zgodnie z założonym kryterium oceny względnej  $y_w$ :

$$y_{wI} = y_B/C$$

Dla oferenta I otrzymujemy na podstawie danych liczbowych:

$$y_{wI} = 0.37/1.25[\text{mln zł}] = 0.296 [1/\text{mln zł}]$$

Dla oferenta II otrzymujemy:

$$y_{wII} = 0.43/1.35[\text{mln zł}] = 0.319 [1/\text{mln zł}]$$

Względna ocena  $y_w$  informuje nas o tym ile korzyści (jakość realizacji prac, czas realizacji prac, dotrzymywanie kosztów, dotrzymywanie czasu realizacji) uzyskuje za każdy 1 milion złotych wydatkowany na danego oferenta.

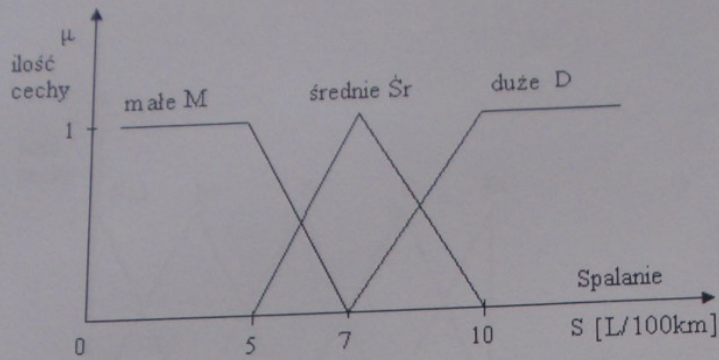
Obliczenia pokazują, że chociaż oferent II żąda więcej pieniędzy za realizację prac, to jednak zgodnie z przyjemnym przez nas kryterium oceny  $y_w$  korzystniej jest zlecić mu prace niż oferentowi I.

Opracowanie powyższej oceny możliwe było dzięki wykorzystaniu lingwistycznej bazy wiedzy, która jest precyzyjnym zapisem naszych poglądów i wymagań co do istotnych cech realizowania robot.

### **Czy tradycyjne, matematyczne kryteria wieloargumentowe mogą konkurować z kryteriami optymalizacyjnymi podanymi w formie lingwistycznej bazy wiedzy?**

Założmy, że zamierzamy kupić używany samochód. Założmy dla uproszczenia problemu, że istotna są dla nas jedynie dwie cechy samochodu. **Najważniejszą cechą jest dla nas spalanie  $S$  [L/100km].** Druga istotna cecha jest ilość km przejechanych przez samochód: **przebieg  $P$  [tys.km].**

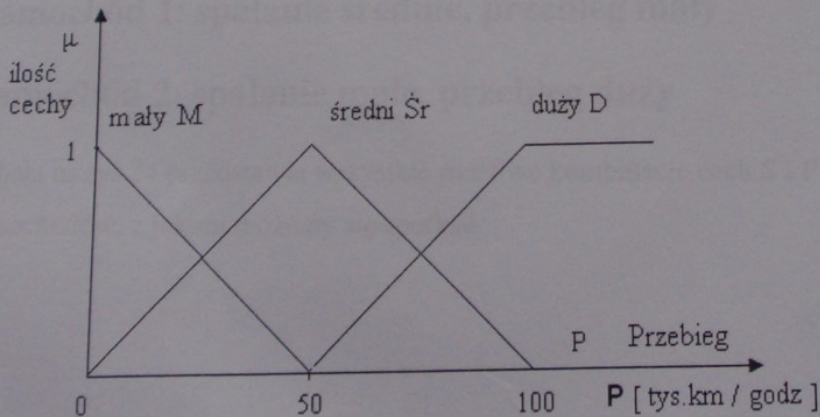
Założmy że według naszych poglądów spalanie na szosie poniżej **5 [L/100km]** uważamy za **małe**, powyżej **10 [L/100km]** za **duże**, a za **typowe średnie** spalanie na szosie uznajemy **7 [L/100km]**. Daje to funkcje cechy ( definicje kwantyfikatorów lingwistycznych) jak na rys.21



**Rys. 21** Definicja małego, średniego i dużego spalania paliwa przez samochód wg. naszych poglądów.

Założmy, że według naszych poglądów za **duży przebieg** samochodu uważamy  $P \geq 100$  tys.km, za **typowy średni** 50 tys.km, a za **typowy mały** 0 km.

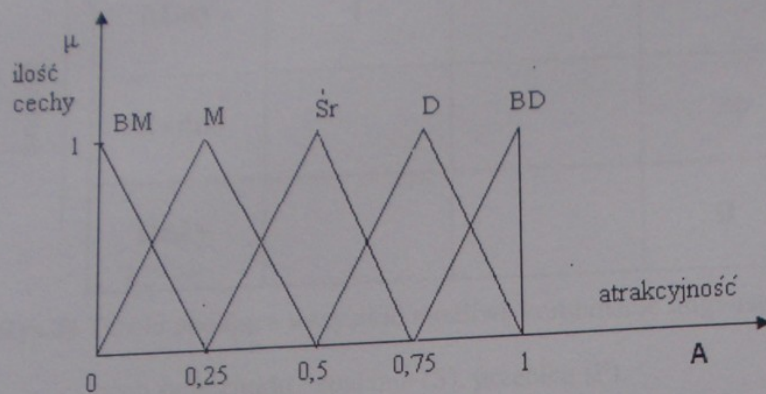
Wyrażeniem powyższych poglądów są funkcje cechy przedstawione na rys. 22.



**Rys.22** Definicje lingwistycznych kwantyfikatorów mały, średni, duży przebieg wynikający z naszych poglądów.

Założmy że do oceny atrakcyjności A samochodu wynikającej ze spalania i przebiegu stosować będziemy 5 kwantyfikatorów i bardzo mała, mała, średnia,

duża i bardzo duża atrakcyjność. Rysunek 23 przedstawia funkcje cechy tych kwantyfikatorów.



Rys.23 Definicje lingwistycznych kwantyfikatorów atrakcyjności samochodu bardzo mała, mała, średnia, duża i bardzo duża atrakcyjność.

Samochody, które spotykamy na giełdzie w ogłoszeniach gazetowych i internetowych, itd., mogą mieć różne wartości cech, np.

**Samochód 1: spalanie średnie, przebieg mały**

**Samochód 2: spalanie małe, przebieg duży**

Tabela na rys.24 przedstawia wszystkie możliwe kombinacje cech S i P samochodów, z jakimi możemy się spotkać.

		← Spalanie			
		[L/100km]	Małe	Średnie	Duże
↑ Przebieg	P	S			
	[ tys. km ]				
	Mały		1		
	Średni				
	Duży				9

Rys.24 Tabela podająca wszystkie możliwe kombinacje lingwistycznych wartości cech samochodu : spalanie (S), przebieg (P).

Do przyporządkowania lingwistycznych wartości atrakcyjności samochodu (BM, M, SR, DB, D) poszczególnym kombinacjom wartości cech S i P zastosujemy **metodę rankingową**.

Bez wątplenia **pierwsze miejsce w rankingu** atrakcyjności zajmie samochód o wartościach cech:

**(spalanie małe) AND (przebieg mały): ranking = 1**

Bez wątplenia **najgorszym samochodem będzie** samochód o wartościach cech:

**(spalanie duże) AND (przebieg duży): ranking = 9**

Te niebudzące wątpliwości pozycje rankingowe podano w tabeli na rys. 24.

Teraz należy ustalić pozycję rankingową pozostałych kombinacji wartości cech.

Rozpatrzmy na początek wszystkie kombinacje sąsiadujące z kombinacją o rankingu 1 na rys. 27.

**Samochód A: (spalanie małe) AND (przebieg średni)**

**Samochód B: (spalanie średnie) AND (przebieg mały)**

**Samochód C: (spalanie średnie) AND (przebieg średni)**

Bez wątpienia samochód C ma najmniej korzystne wartości cech spośród trzech powyższych samochodów. Wobec tego najlepszym może być albo samochód A albo B. Na pierwszy rzut oka obydwa samochody są równoważne. Jednak pamiętając o tym, że **małe spalanie jest dla nas bardziej istotne niż mały przebieg** (nasze osobiste preferencje) najwyższą pozycję rankingową przyznajemy samochodowi A, kolejną B, a ostatnią C, rys. 25.

S [l/100km] \ P [tyś.km]	Małe	Średnie	Duże
Mały	1	3	
Średni	2	4	
Duży			9

Rys. 25. Częściowo wypełniona tabela rankingowa.

Rozpatrzmy teraz 2 kombinacje wartości cech sąsiadujące z rankingowo najniższą kombinacją 9 na rys. 25.

**Samochód E: (spalanie średnie) AND (przebieg duży)**

**Samochód F: (spalanie duże) AND (przebieg średni)**

Obydwie kombinacje pozornie są równoważne. Ponieważ jednak **spalanie jest dla nas istotniejsze niż przebieg** to wyższą pozycję w rankingu musi zająć samochód E a niższą F, rys. 26.

S [l/100km] \ P [tyś.km]	Małe	Średnie	Duże
Mały	1	3	
Średni	2	4	8
Duży		7	9

Rys. 26. Częściowo wypełniona tabela rankingowa kombinacji wartości cech samochodu.

Do rozpatrzenia pozostają jeszcze 2 kombinacje wartości cech:

**Samochód G: (spalanie małe) AND (przebieg duży)**

**Samochód H: (spalanie duże) AND (przebieg mały)**

Jeżeli **spalanie jest dla nas istotniejsze niż przebieg**, to samochód G musi mieć wyższą pozycję rankingową (5) niż samochód H(6). W ten sposób uzupełniamy tabelę rankingową, rys. 27.



S [l/100km] \ P [tyś.km]	Małe	Średnie	Duże
Mały	1	3	6
Średni	2	4	8
Duży	5	7	9

Rys. 27. Pełna tabela rankingowa kombinacji wartości cech samochodu.

**Uwaga!**

Ustalanie kolejności rankingowej nie zawsze jest łatwe i oczywiste. **Trudności** pojawiają się wówczas, gdy nasze **preferencje nie są wyraźne** (np. nie bardzo wiemy czy ważniejsze jest dla nas spalanie czy przebieg). Trudności pojawiają się też wówczas, gdy sformujemy **więcej kwantyfikatorów lingwistycznych** dla poszczególnych cech np.

**5 kwantyfikatorów: ( BM, M, Śr, D, BD)**

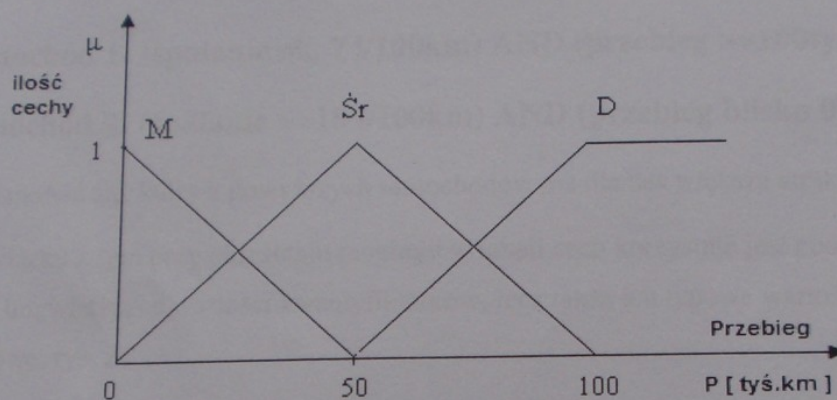
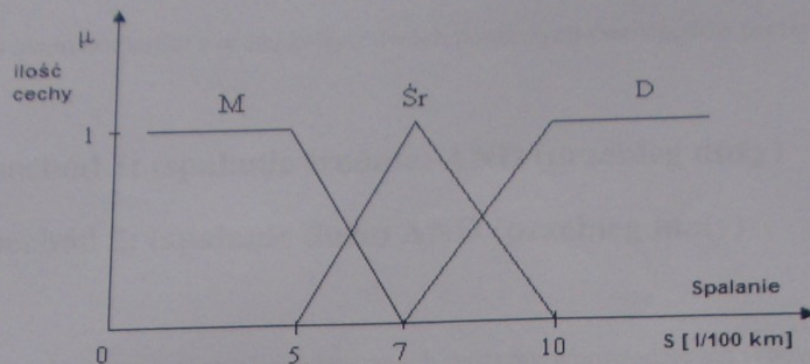
**7 kwantyfikatorów: ( MIN, BM, M, Śr, D, BD, MAX)**

**9 kwantyfikatorów etc.**

Wówczas różnica między poszczególnymi kombinacjami cech jest coraz mniejsza i trudniejsza do wykrycia.

W takiej sytuacji bardzo pomocna może być **metoda „typowych wartości kwantyfikatorów lingwistycznych”**.

Na rys.28 pokazane są funkcje cechy poszczególnych kwantyfikatorów.



Rys. 28. Funkcje cechy lingwistycznych kwantyfikatorów spalania i przebiegu samochodu stosowane w rozpatrywanym problemie.

Typowymi wartościami kwantyfikatorów spalania są:

$$TWS (M) = 5, \quad TWS (\acute{S}r) = 7, \quad TWS (D) = 10$$

Typowymi wartościami kwantyfikatorów przebiegu są:

$$TWP (M) = 0, \quad TWP (\acute{S}r) = 50, \quad TWP (D) = 100$$

Jeżeli mamy trudności z oceną który z dwóch poniższych samochodów preferujemy:

**Samochód 1: (spalanie średnie) AND (przebieg duży)**

**Samochód 2: (spalanie duże) AND (przebieg mały)**

to zamiast kwantyfikatorów lingwistycznych możemy wprowadzić ich typowe wartości liczbowe, jak poniżej.

**Samochód 1: (spalanie ok. 7 l/100km) AND (przebieg  $\geq 100$ tyś. km)**

**Samochód 2: (spalanie  $\geq 10$  l/100km) AND (przebieg blisko 0)**

i zastanowić się, który z powyższych samochodów ma dla nas większą atrakcyjność.

W związku z tym przy określaniu rankingu w tabeli cech korzystnie jest podać nie tylko lingwistyczne wartości kwantyfikatorów, lecz także ich typowe wartości liczbowe, rys. 24.

S [l/100km] \ P [tyś.km]	Małe (około 5)	Średnie (około 7)	Duże (ponad 10)
Mały (blisko 0)	1		
Średni (około 50)			
Duży (ponad 100)			9

**Rys. 24.** Tabela z kwantyfikatorami lingwistycznymi oraz ich typowymi wartościami liczbowymi.

Ponieważ ustalony został już ranking poszczególnych kombinacji wartości cech samochodu, rys.27 można teraz przystąpić do przyporządkowania lingwistycznych wartości atrakcyjności do poszczególnych kombinacji cech.

Ranki ng	Kombinacja cech	Kwantyfikatory atrakcyjności
1	(małe spalanie) AND (mały przebieg)	bardzo duża (ok. 1)
2	(małe spalanie) AND (średni przebieg)	duża (ok. 0,75)
3	(średnie spalanie) AND (mały przebieg)	średnio-duża (ok. 0,675)
4	(średnie spalanie) AND (średni przebieg)	średnia (ok. 0,5)
5	(małe spalanie) AND (duży przebieg)	bardzo mała (ok. 0)
6	(duże spalanie) AND (mały przebieg)	zero (dokładnie 0)
7	(średnie spalanie) AND (duży przebieg)	zero (dokładnie 0)
8	(duże spalanie) AND (średni przebieg)	zero (dokładnie 0)
9	(duże spalanie) AND (duży przebieg)	zero (dokładnie 0)

Rys. 25. Tabela przyporządkowań lingwistycznych kwantyfikatorów atrakcyjności samochodu do poszczególnych kombinacji wartości cech.

Tabela z rys. 28 może być przedstawiona w formie pokazanej na rys. 29.

S [l/100km] \ P [tyś.km]	Małe (około 5)	Średnie (około 7)	Duże (ponad 10)
Mały (blisko 0)	<b>BD</b> (ok. 1)	<b>Śr-D</b> (ok. 0,675)	<b>Zero</b> (dokładnie 0)
Średni (około 50)	<b>D</b> (ok. 0,75)	<b>Śr</b> (ok. 0,5)	<b>Zero</b> (dokładnie 0)
Duży (ponad 100)	<b>BM</b> (ok. 0)	<b>Zero</b> (dokładnie 0)	<b>Zero</b> (dokładnie 0)

atrakcyjność

**Rys. 28.** Lingwistyczna baza wiedzy określająca zależność między atrakcyjnością samochodu oraz wartościami jego cech: spalanie i przebieg.

Każde z pól w tabeli z rys. 28 określa jedną regułę lingwistyczną, np:

**IF [(spalanie średnie) AND (przebieg duży) THEN (atrakcyjność zero).**

Określenie **atrakcyjność = zero** oznacza, że samochodu takiego na pewno nie chcielibyśmy kupić, gdyż w ogóle nie spełnia naszych wewnętrznych oczekiwań, nawet w minimalnym stopniu.

#### Uwaga – bazy indywidualne i grupowe

Lingwistyczne BW reprezentuje poglądy tylko tej osoby, która ją utworzyła. Inne osoby mogą mieć inne indywidualne poglądy na atrakcyjność samochodu. Można jednak skonstruować **przeciętną bazę wiedzy** reprezentującą poglądy wielu osób (grupy osób). Baza taka będzie miała bardziej uniwersalny charakter niż baza indywidualna.

Na podstawie tabeli z rys. 28 zawierającej typowe liczbowe wartości poszczególnych kwantyfikatorów można narysować powierzchnię zależności

**Atrakcyjność = f (spalanie, przebieg).**

Przykładową regułę lingwistyczną:

**IF (spalanie małe) AND (przebieg średni) THEN (atrakcyjność duża)**

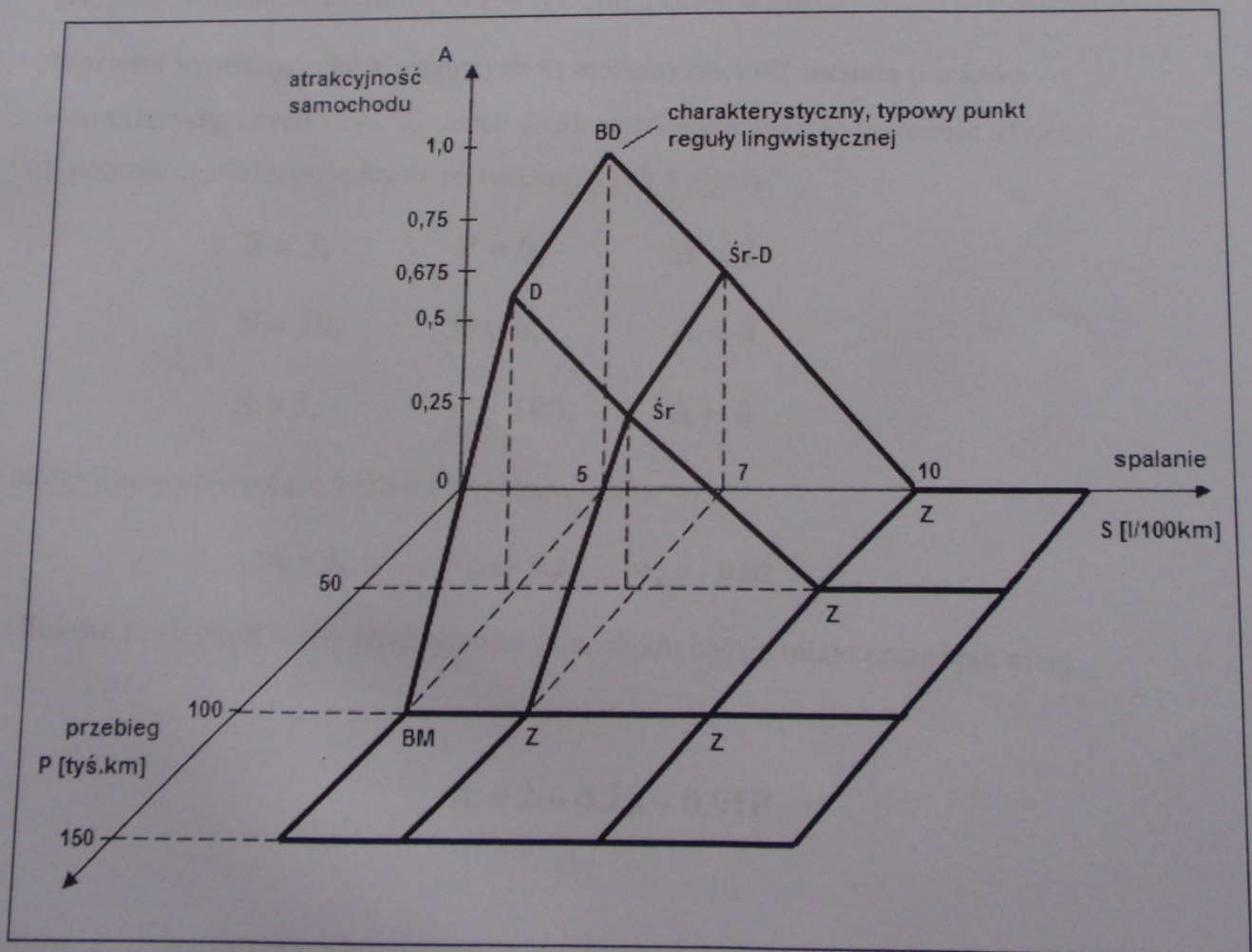
można zapisać w formie:

**IF (spalanie ok. 5) AND (przebieg ok. 50) THEN (atrakcyjność ok. 0,75).**

Typowymi wartościami cech podanych w powyższej regule odpowiada jeden punkt o współrzędnych:

$$S = 5 \quad P = 50 \quad A = 0.75$$

Podobnie każda z pozostałych ośmiu reguł tworzy jeden punkt w przestrzeni cech. Nanieśnienie wszystkich punktów i ich połączenie umożliwia uzyskanie powierzchni modelującej rozpatrywaną zależność, rys. 29.



Rys. 29. Powierzchnia zależności atrakcyjność samochodu = f(spalanie, przebieg)

Uwaga!

Zauważ, że **powierzchnia** zależności  $A = f(S, P)$  na rys. 29 **nie jest płaska**, tzn. nie jest liniowa. Jest zakrzywiona (nieliniowa) i z tego względu nie można jej przybliżyć równaniem liniowym.

$$A = w_0 + w_1 S + w_2 P$$

w którym  $w_0, w_1, w_2$  są współczynnikami wagowymi i jednocześnie przeliczeniowymi. Z jednej strony **współczynniki  $w_1, w_2$  muszą odzwierciedlać nasze preferencje** względem spalania i przebiegu samochodu. Z drugiej strony **muszą umożliwić przeliczenie spalania i przebiegu na atrakcyjność samochodu**.

Zauważ bowiem, że atrakcyjność  $A$  przyjmuje wartości w zakresie  $[0, 1]$ . Natomiast spalanie może mieć wartości w zakresie  $[5, 13$  (l/100km)] a nawet więcej, przebieg przyjmuje wartości w zakresie  $[0, 150$  tys. km] a nawet więcej.

**Wartości współczynników wagowych  $w_i$  możemy określić metodą punktów charakterystycznych** z rys. 29. Jeżeli do identyfikacji liniowego kryterium użyjemy 3 punktów charakterystycznych, reprezentujących 3 reguły,

$$S = 5, \quad P = 0, \quad A = 1$$

$$S = 10, \quad P = 0, \quad A = 0$$

$$S = 5, \quad P = 100, \quad A = 0$$

to uzyskamy następujące wartości współczynników  $w_i$ :

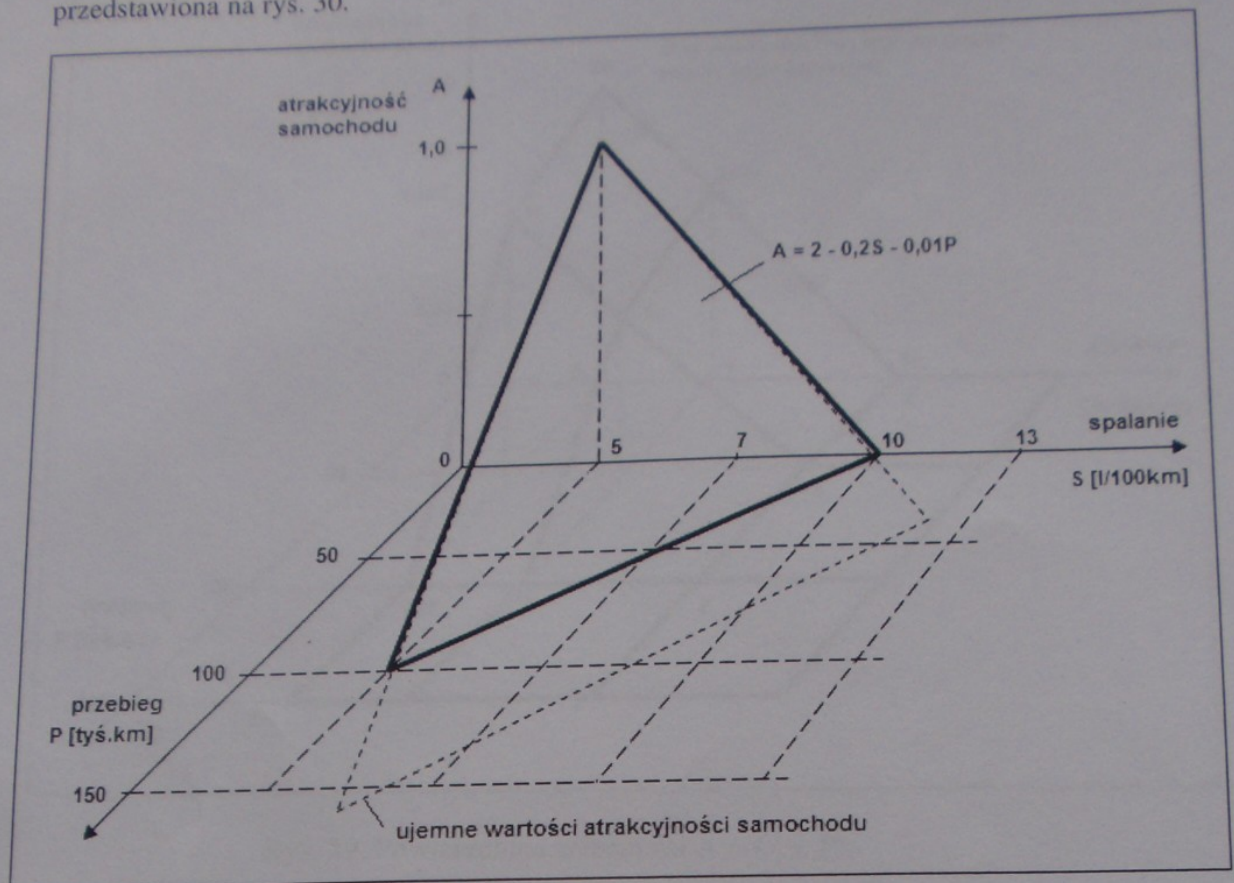
$$w_0 = 2, \quad w_1 = -0,2, \quad w_2 = -0,01$$

a **liniowe kryterium** oceny atrakcyjności samochodu będzie miało postać jak niżej:

$$A = 2 - 0,2S - 0,01P$$

Zauważ, że ze wzrostem zarówno spalania jak i przebiegu maleje atrakcyjność samochodu (**współczynniki wagowe są ujemne**).

Powyższej, liniowej formie kryterium odpowiada powierzchnia kryterialna przedstawiona na rys. 30.

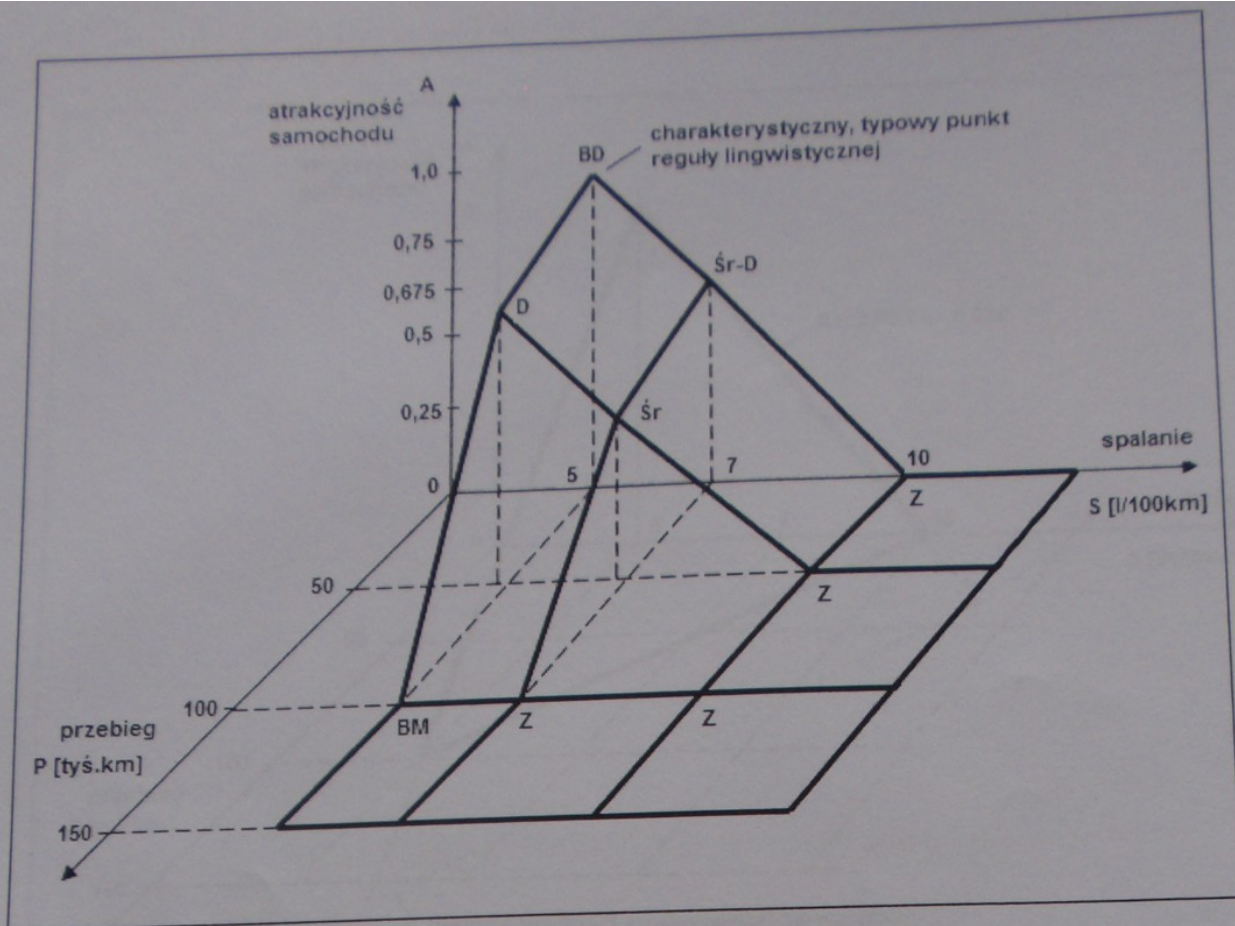


Rys. 30. Powierzchnia liniowego kryterium oceny atrakcyjności samochodu na podstawie wartości spalania i przebiegu.

Zauważ, że dla np.  $S = 8$  [l/100km] i  $P = 75$  [tyś. km] stosując wzór  $A = 2 - 0,2S - 0,01P$  uzyskujemy **ujemną atrakcyjność**  $A = - 0,35(?)$ . Tymczasem atrakcyjność powinna, zgodnie z założeniem, zawierać się w zakresie [ 0, 1].

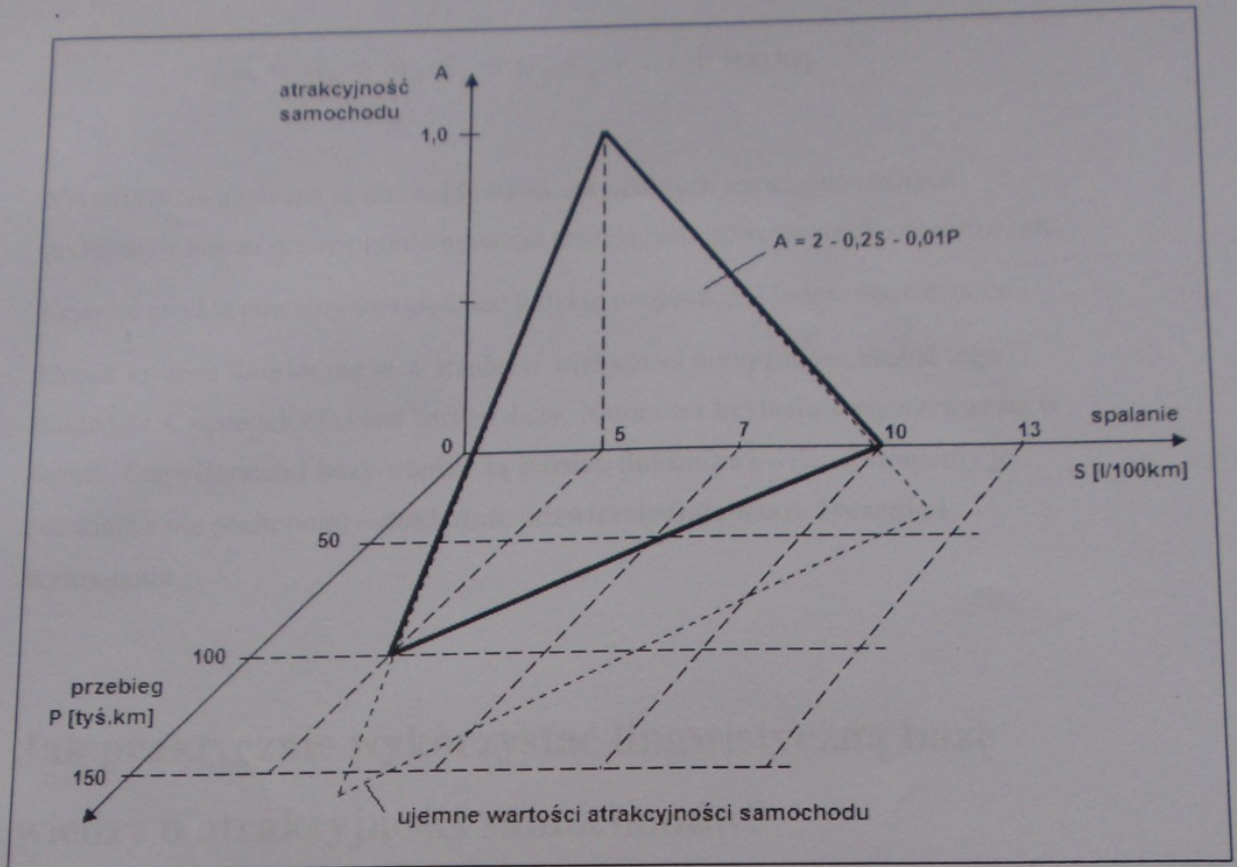
Porównaj teraz powierzchnie atrakcyjności samochodu wynikającą z lingwistycznej bazy wiedzy , rys.29, oraz z kryterium liniowego, rys. 30.





Rys. 29. Powierzchnia zależności  $A = f(S, P)$ :

atrakcyjność samochodu =  $f(\text{spalanie}, \text{przebieg})$



Rys. 30. Powierzchnia liniowego kryterium oceny atrakcyjności samochodu na podstawie wartości spalania i przebiegu.

Porównanie rys. 29 i 30 wyraźnie pokazuje, że **żadna płaska powierzchnia kryterialna** określona przez liniowy model atrakcyjności samochodu, rys. 30

$$A = w_0 + w_1S + w_2P$$

nie jest w stanie oddać rzeczywistych preferencji człowieka określonych powierzchnią na rys. 29. Jakikolwiek wartości  $w_0$ ,  $w_1$ ,  $w_2$  dobralibyśmy w liniowym modelu atrakcyjności zawsze otrzymamy płaską powierzchnię odbiegającą od rzeczywistej powierzchni atrakcyjności dokładnie odpowiadającej naszej prawdziwej, lingwistycznej bazie wiedzy.

To spostrzeżenie ma aspekt szerszy, wychodzący poza omawiany przykład. W nauce i technice powszechnie stosuje się liniowe kryteria jakości typu:

$$K = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n$$

Kryteria takie używane są do znajdowania optymalnych rozwiązań różnych problemów lub do oceny i porównywania istniejących już rozwiązań, obiektów, etc.

Kryteria zwykle powinny uwzględniać ludzkie preferencje i ludzki sposób oceny.

Jednak kryteria liniowe nie są w stanie, w większości przypadków, zrobić tego dokładnie. Często ich błąd jest bardzo duży. Natomiast **kryteria wygenerowane w formie lingwistycznej bazy wiedzy są bardzo dokładne** (o ile sformujemy je starannie a nie pochopnie) – **dokładnie odzwierciedlają nasze życzenia i wymagania.**

## **Jak praktycznie wykorzystać lingwistyczną bazę wiedzy o atrakcyjności samochodów?**

Założmy, że chcesz obliczyć atrakcyjność samochodu (zgodnie z twymi poglądami) o następujących wartościach cech:

**spalanie = 8,3 [l/100km], przebieg = 38 [tyś. km]**

Dane te wprowadzasz do programu komputerowego zawierającego Twoją lingwistyczną bazę wiedzy, który oblicza konkretną liczbową wartość atrakcyjności z zakresu  $A \in [0, 1]$ , gdzie:

**1 – atrakcyjność maksymalna, 0 – minimalna.**

## Sprawdź czy zrozumiałeś!

1. Zastanów się jakie cechy, wg Ciebie, są najważniejsze w ocenie posiadłości (dom z ogrodem). Ułóż lingwistyczną bazę wiedzy, zastosuj metodę dekompozycji.
2. Spalanie paliwa przez samochód zależy od szybkości jazdy i kąta pochylenia drogi. Ułóż lingwistyczną bazę wiedzy o tej zależności.
3. Wytrzymałość stali na zrywanie rośnie z zawartością niklu, maleje z zawartością krzemu. Ułóż lingwistyczną bazę wiedzy o tej zależności.
4. Smak kawy zależy od ilości wyspanego cukru i śmietanki. Ułóż lingwistyczną bazę wiedzy o tej zależności.
5. Znajdź problem, w którym jedna wielkość zależy od innych i ułóż lingwistyczną bazę wiedzy o tej zależności. W razie potrzeby zastosuj dekompozycje.