Wstęp

1982 – prof. Pawlak sformułował nową teorię:

Teorię zbiorów przybliżonych

(ang. rough set theory)

stanowiącą rozwinięcie klasycznej teorii zbiorów.

Głównym celem jest dostarczanie narzędzi dla problemu aproksymacji pojęć (zbiorów).

Zastosowania w systemach decyzyjnych:

- Redukcja danych, selekcja ważnych atrybutów;
- Generowanie reguł decyzyjnych;
- Odkrywanie wzorców z danych: szablony, reguły asocjacyjne;
- Odkrywanie zależności w danych.

Podstawowe pojęcia

Zbiór przybliżony (ang. rough set) – obiekt matematyczny zbudowany w oparciu o logikę trójwartościową.

Universum – zebrany, skończony, niepusty zbiór informacji.

Podstawowe pojęcia cd.

W zbiorze **U(universum)** wyróżnia się:

- Elementy zbioru **U** zwane **obiektami.**
- Wszystkie podzbiory $X \subseteq U$

  są **„pojęciami”** w U
- Dowolna rodzina pojęć zbioru **U** to wiedza o **U**

Podstawowe pojęcia cd.

W zbiorze **U(universum)** wyróżnia się:

- każda rodzina rozłącznych i niepustych zbiorów **U**, spełniająca warunek iż ich suma jest równa całemu zbiorowi **U**, nazywa się **podziałem** (klasyfikacją) zbioru **U**,
- para $\mathbf{K = (U,C)}$, gdzie $\mathbf{C = \{C_1, C_2, \ldots, C_n\}}$ jest dowolną rodziną podziałów zbioru **U**, określa się mianem **bazy wiedzy** o **U**.

Podstawowe pojęcia cd.

Ostatni punkt jest równoważny poniższemu pokazanemu w ujęciu relacyjnym:

- Zakładając, że **U** to **uniwersum**, to para

  K = (U,R), gdzie **R = {R_1, R_2, ..., R_n}** jest dowolną rodziną relacji równoważności określonych na **U**, nazywa się bazą wiedzy o **U**.

Podstawowe pojęcia cd.

Pojęcie podstawowe:

Zakładając iż **U** jest uniwersum i **K = (U,R)** jest bazą wiedzy o **U**, to każdą klasę abstrakcji dowolnej relacji R $\in$ **R** nazywamy pojęciem podstawowym w bazie wiedzy **K**.

Pojęcie podstawowe jest na ten moment nośnikiem informacji podstawowej o własnościach obiektów danego uniwersum.

Podstawowe pojęcia cd.

Pojęcie elementarne:

Zakładając iż **U** to uniwersum i **K = (U,R)** jest bazą wiedzy o **U** oraz $\cap R$ jest relacją $R_1 \cap \ldots \cap R_n$,

gdzie $R_i \in R$ dla $i \in \{1,2, \ldots, n\}$, to każdą klasę abstrakcji relacji $\cap R$ nazywa się pojęciem elementarnym w bazie wiedzy **K**.

Pojęcie elementarne, zwane także pojęciem atomowym to elementarne – najmniejsze pojęcie, za pomocą którego możemy opisać nasze uniwersum.

System informacyjny

System informacyjny (ang. Information system) – uporządkowana para **A = (U,A)**, gdzie:

- **U** jest skończonym, niepustym zbiorem, zwanym **uniwersum**, przy czym elementy zbioru **U** nazywamy **obiektami.**
- **A** jest skończonym, niepustym zbiorem atrybutów (własności, cech), gdzie każdy atrybut a $\in$ A jest funkcją

  $\boldsymbol{a: U \rightarrow V_a}$,

  przy czym $\mathbf{V_a}$, jest zbiorem wartości atrybutu.

Dla B $\subseteq$ ***A*** definiujemy

- sygnatura obiektu (ang. -*information vector*) jako

$$\inf_{\boldsymbol{B}}(x)=\{(a,a(x)):a \in B\}$$

- Zbiór sygnatur względem o obiektach z (ang. -*information set*):

$$INS(U)=\{\inf_{\boldsymbol{B}}(x) : x \in U \}$$

System informacyjny (przykład)

	Ból głowy	Temperatura [°C]
x_1	Tak	36-37
x_2	Nie	36-37
x_3	Tak	37-38
x_4	Nie	36-37
x_5	Tak	37-38
x_6	Tak	38-39

System informacyjny (przykład 2)

	Ból głowy	Bóle mięśni	Dreszcze	Temperatura [°C]
x_1	Tak	Nie	Nie	36-37
x_2	Nie	Nie	Nie	36-37
x_3	Tak	Tak	Nie	37-38
x_4	Tak	Nie	Nie	36-37
x_5	Tak	Tak	Tak	37-38
x_6	Tak	Tak	Tak	38-39

System decyzyjny

System decyzyjny (ang. decision system) jest to system informacyjny z dodatkowym atrybutem ***d***, nazywanym atrybutem decyzyjnym (decision attribute).

System decyzyjny możemy również zapisać w postaci funkcji: $\boldsymbol{D = (U, A \cup d)}$, gdzie $\boldsymbol{d \notin A}$, jest **atrybutem decyzyjnym**. **Atrybut decyzyjny** może przyjmować wiele wartości, ale jest to najczęściej wartość binarna (prawda albo fałsz). Pozostałe atrybuty $\boldsymbol{a \in A - d}$, nazywamy **atrybutami warunkowymi** (conditional attributes).

Tablica decyzyjna powstaje ze zwykłych tablic danych poprzez sprecyzowanie:

- Atrybutów (nazwanych warunkowymi): cechy, których wartości na obiektach są dostępne, np. pomiary, parametry, dane osobowe, …
- Decyzji (atrybut decyzyjny):, t.j. cecha “ukryta” związana z pewną znaną częściowo wiedzą o pewnym pojęciu

- Decyzja jest znana tylko dla obiektów z (treningowej) tablicy decyzyjnej;
- Jest podana przez eksperta (np. lekarza) lub na podstawie późniejszych obserwacji (np. ocena giełdy);
- Chcemy podać metodę jej wyznaczania dla dowolnych obiektów na podstawie wartości atrybutów warunkowych na tych obiektach.

System decyzyjny (przykład)

	Ból głowy	**Temperatura [°C]**	***Grypa***
$\mathbf{x_1}$	**Tak**	**36-37**	***Nie***
$\mathbf{x_2}$	**Nie**	**36-37**	***Nie***
$\mathbf{x_3}$	**Tak**	**37-38**	***Tak***
$\mathbf{x_4}$	**Nie**	**36-37**	***Nie***
$\mathbf{x_5}$	**Tak**	**37-38**	***Nie***
$\mathbf{x_6}$	**Tak**	**38-39**	***Tak***

System decyzyjny (przykład 2)

	Ból głowy	**Bóle mięśni**	**Dreszcze**	**Temperatura [°C]**	***Grypa***
x_1	**Tak**	**Nie**	**Nie**	**36-37**	***Nie***
x_2	**Nie**	**Nie**	**Nie**	**36-37**	***Nie***
x_3	**Tak**	**Tak**	**Nie**	**37-38**	***Tak***
x_4	**Tak**	**Nie**	**Nie**	**36-37**	***Nie***
x_5	**Tak**	**Tak**	**Tak**	**37-38**	***Nie***
x_6	**Tak**	**Tak**	**Tak**	**38-39**	***Tak***

Nierozróżnialność

Przypomnienie:

Binarna relacja $\boldsymbol{R} \subseteq \boldsymbol{X} \times \boldsymbol{X}$, która jest **zwrotna** (obiekt jest w relacji z samym sobą, tzn. *xRx*), **symetryczna** (jeżeli *xRy* to *yRx*) oraz **przechodnia** (jeśli *xRy* i *yRz* to *xRz*) jest nazywana **relacją równoważności**.

Relacja równoważności na danym zbiorze wyznacza klasy równoważności, czyli takie zbiory elementów, że każde dwa elementy są ze sobą w relacji.

Nierozróżnialność

- Niech dany będzie system informacyjny

  $\boldsymbol{S = (U, A)}$ i zbiór $\boldsymbol{B}$, taki że, $\boldsymbol{B \subset A}$. Przez $\boldsymbol{IND_A(B)}$ oznaczamy relację równoważności określoną następująco:

  $$\boldsymbol{IND_A(B) = \{(x, y) \in U^2 : \forall a \in B \ \ a(x)=a(y)\}}$$

Relację tą nazywamy relacją **nierozróżnialności (ang. indiscernibility)**.

Dane są obiekty x,y $\in$ U i zbiór atrybutów B$\subseteq$***A***, mówimy, że

- x,y są rozróżnialne przez wtw, gdy istnieje a $\in B$ taki, że a(x) $\neq$ a(y);
- x,y są nierozróżnialne przez B, jeśli one są identyczne na B , tzn. a(x)=a(y) dla każdego a $\in B$;
- $[x]_{\boldsymbol{B}}$ = zbiór obiektów nierozróżnialnych z x przez B.

Dla każdych obiektów :

- albo $[x]_{\boldsymbol{B}} = [y]_{\boldsymbol{B}}$;
- albo $[x]_{\boldsymbol{B}} \cap [y]_{\boldsymbol{B}} = \varnothing$.

Każdy zbiór atrybutów B wyznacza podział zbioru obiektów na klasy nierozróżnialności.

Nierozróżnialność (przykład)

IND({Ból głowy}) = $\{\{x_1, x_3, x_5, x_6\}, \{x_2, x_4\}\}$

IND({Temp.}) = $\{\{x_1, x_2, x_4\},\{x_3, x_5\},\{x_6\}\}$

IND({Ból głowy, Temp.}) = $\{\{x_1\},\{x_2, x_4\},\{x_3, x_5\},\{x_6\}\}$

	Ból głowy	Temperatura [°C]	*Grypa*
x_1	Tak	36-37	*Nie*
x_2	Nie	36-37	*Nie*
x_3	Tak	37-38	*Tak*
x_4	Nie	36-37	*Nie*
x_5	Tak	37-38	*Nie*
x_6	Tak	38-39	*Tak*

Nierozróżnialność (przykład 2)

IND({Ból głowy}) = {{x_1, x_3, x_4, x_5, x_6}, {x_2}}
IND({Bóle mięśni}) ={{x_3,x_5 ,x_6},{x_1, x_2, x_4}}
IND({Dreszcze}) = {{x_1, x_2, x_3, x_4}, {x_5, x_6}}
IND(Temperatura) = {{x_1, x_2 , x_4}, {x_3, x_5}, {x_6}}
IND({Ból głowy, Bóle mięśni, Dreszcze, Temperatura}) = {{x_1, x_4}, {x_2}, {x_3}, {x_5}, {x_6}}

	Ból głowy	**Bóle mięśni**	**Dreszcze**	**Temperatura [°C]**	***Grypa***
x_1	**Tak**	**Nie**	**Nie**	**36-37**	***Nie***
x_2	**Nie**	**Nie**	**Nie**	**36-37**	***Nie***
x_3	**Tak**	**Tak**	**Nie**	**37-38**	***Tak***
x_4	**Tak**	**Nie**	**Nie**	**36-37**	***Nie***
x_5	**Tak**	**Tak**	**Tak**	**37-38**	***Nie***
x_6	**Tak**	**Tak**	**Tak**	**38-39**	***Tak***

Nierozróżnialność (własności)

- $IND_A(B)$ jest relacją równoważnościową,
- $IND_A(\emptyset) = U \times U$
- $\forall B_1, B_2 \in A$:

$$IND_A(B_1 \cup B_2) = IND_A(B_1) \cap INDA(B_2),$$

- $IND_A(B) = \cap\ IND_A(\{a\})$

Wynikają z definicji relacji nierozróżnialności oraz z podstawowych zasad logiki i teorii mnogości

Nierozróznialność cd.

Relacja nierozróżnialności dzieli nam zbiór obiektów ze względu na zbiór atrybutów ***B*** na klasy równoważności, które oznaczamy przez $[x]_B$.

O ile nie powoduje to dwuznaczności, to zamiast $IND_A(B)$ pisze się $IND(B)$.

Przybliżenie zbioru

Definiujemy ***B*-dolne** przybliżenie zbioru ***X*** przez:

$$\underline{B}X = \{x \mid [x]_B \subseteq X\}$$

B*-górne** przybliżenie zbioru ***X przez:

$$BX = \{x \mid [x]_B \cap X \neq \emptyset\}$$

Obiekty należące do **dolnego przybliżenie** zbioru ***X***, **na pewno** należą do ***X***. Obiekty należące do **górnego przybliżenia** zbioru ***X***, **możliwe że,** należą do zbioru ***X***.

Każdy zbiór obiektów X (np. klasa decyzyjna, pojęcie) może być opisany za pomocą atrybutów ze zbioru dokładnie lub w przybliżeniu

- dokładny opis: jeśli X jest sumą pewnych klas nierozróżnialności definiowanych przez (ZBIORY DOKŁADNE)
- przybliżony opis: w przeciwnym przypadku (ZBIORY PRZYBLIŻONE)

Przybliżanie zbioru

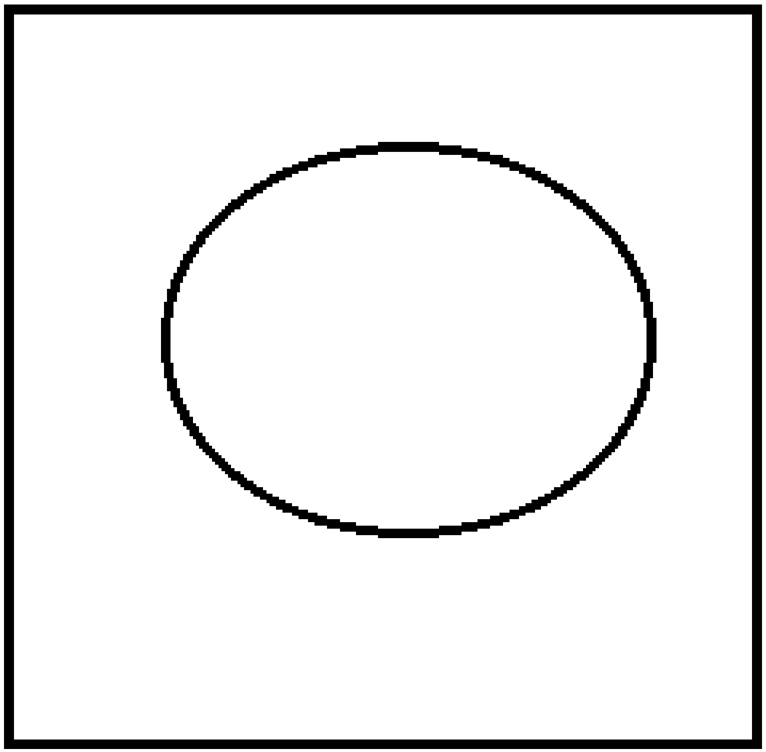
Zbiór $\boldsymbol{BN_B(X) = BX - \underline{B}X}$ nazywamy ***B*-regionem granicznym** (brzegiem) zbioru ***X***. Jest to zbiór tych elementów, których **nie jesteśmy pewni** czy należą do ***X***.

Zbiór $\boldsymbol{U - BX}$ nazywamy obszarem **negatywnym zbioru *X***. Jest to zbiór tych elementów, które **z pewnością** nie należą do ***X***.

- Obszar brzegowy (ang.B -boundary region) pojęcia X zawiera obiekty, dla których nie możemy jednoznacznie zdecydować czy należą one do czy nie do X na podstawie atrybutów z B
- Obszar wewnętrzny (ang. B-inside region of X) zawiera obiekty, które możemy pewnie klasyfikować jako elementy pojęcia X mając do dyspozycji atrybuty z B.
- Zbiór jest przybliżony (ang. rough set) jeśli obszar brzegowy jest niepusty, w przeciwnym przypadku zbiór jest nazwany dokładny (ang. crisp set).

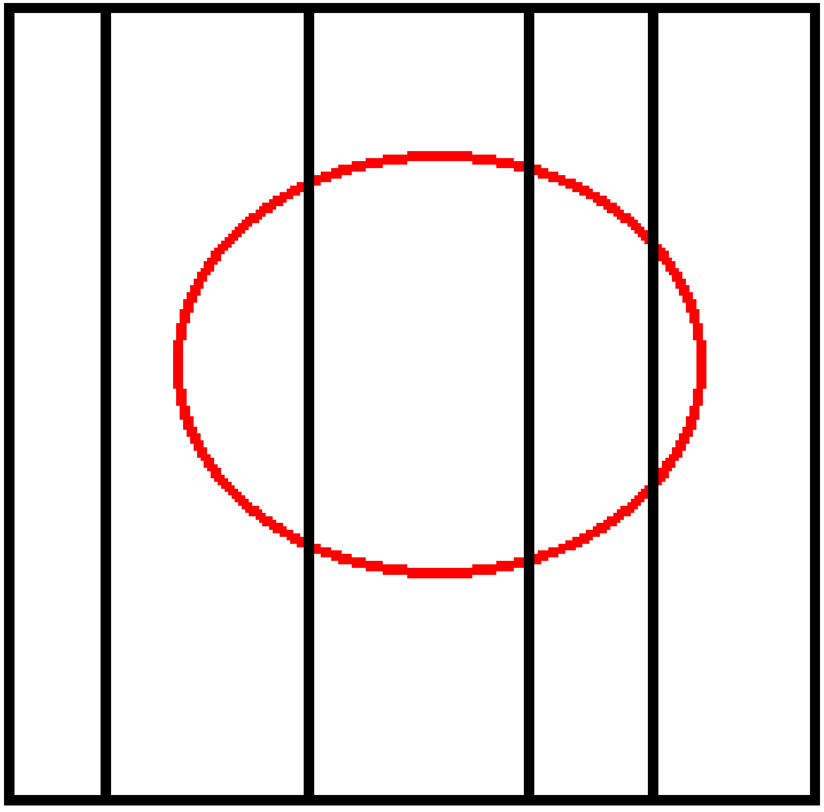
Podsumowanie

Zbiór do przybliżenia, brak użytych atrybutów



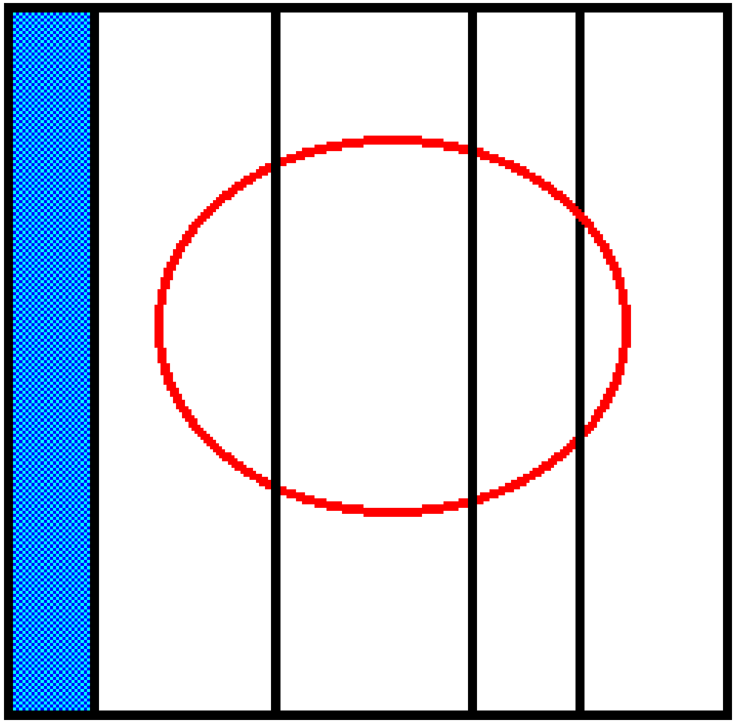
Podsumowanie

Podział z pierwszym atrybutem



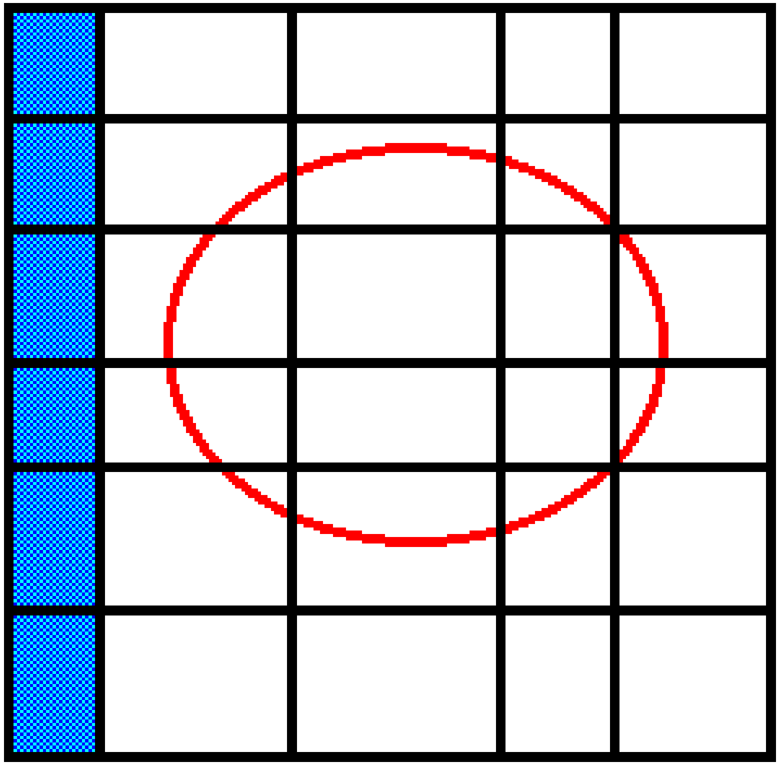
Podsumowanie

Obszar negatywny



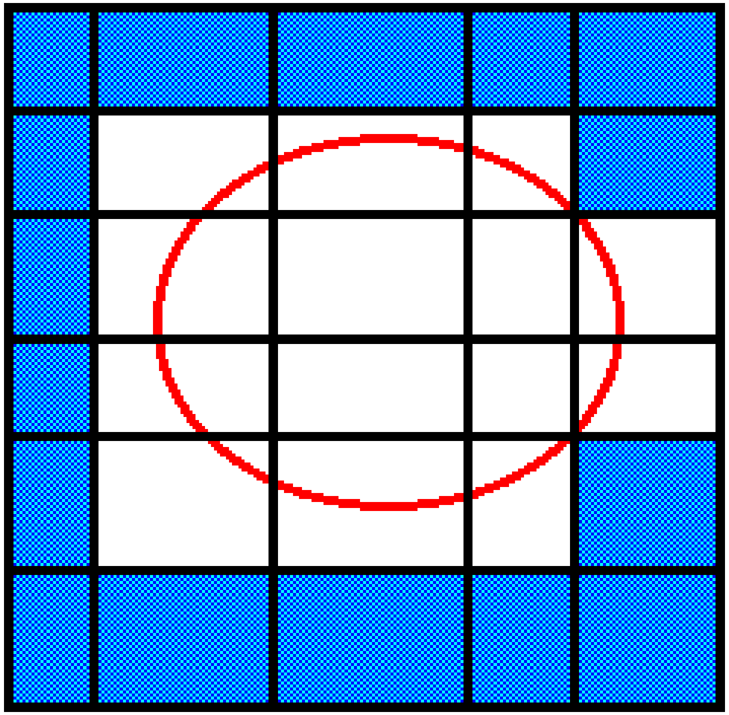
Podsumowanie

Podział z dwoma atrybutami



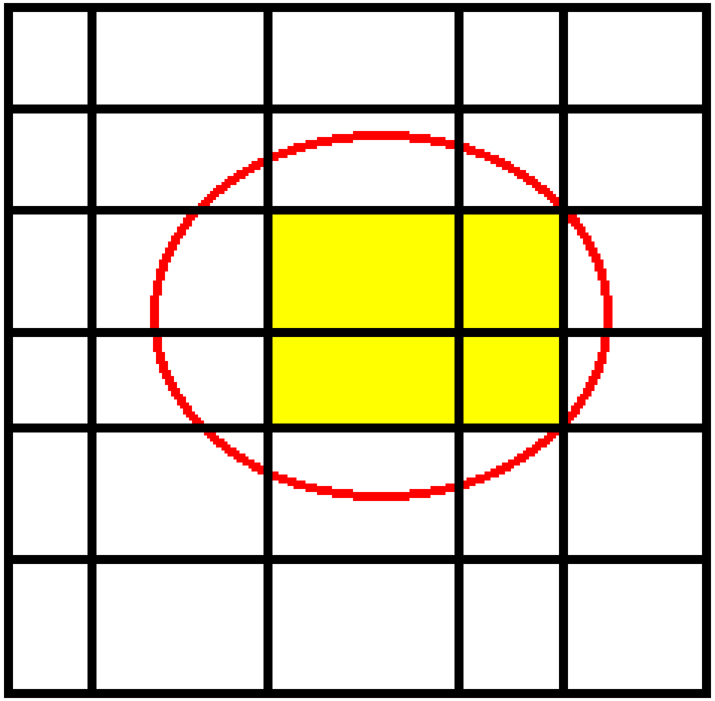
Podsumowanie

Obszar negatywny



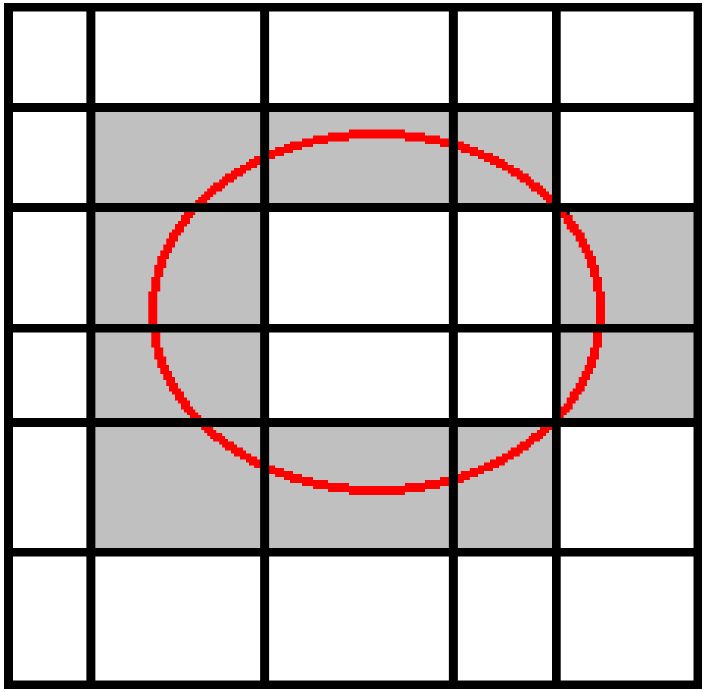
Podsumowanie

Obszar pozytywny - dolne przybliżenie



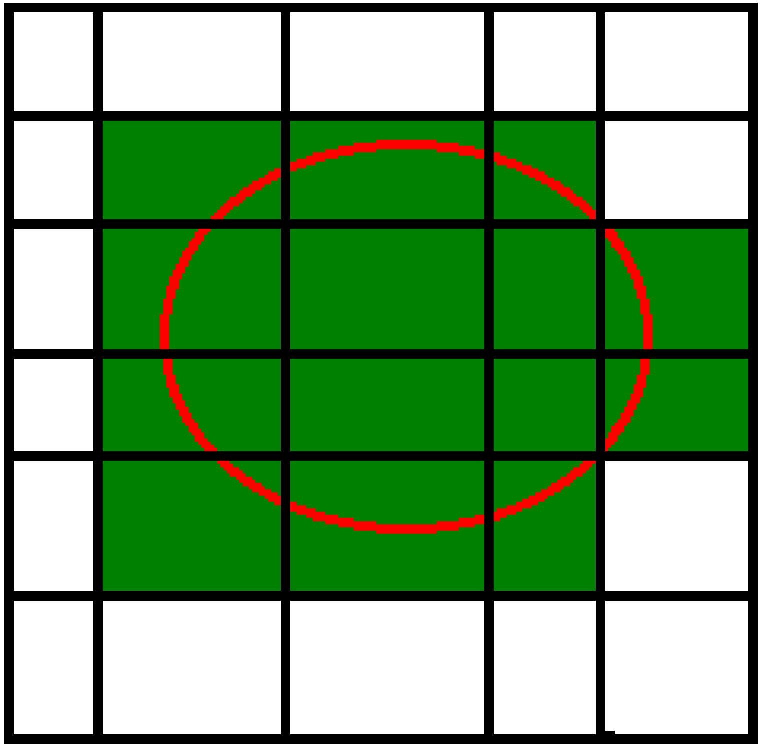
Podsumowanie

Obszar graniczny



Podsumowanie

Górne przybliżenie



Przybliżanie zbioru (przykład)

	Ból głowy	Temperatura [°C]	Grypa
x_1	Tak	36-37	Nie
x_2	Nie	36-37	Nie
x_3	Tak	37-38	Tak
x_4	Nie	36-37	Nie
x_5	Tak	37-38	Nie
x_6	Tak	38-39	Tak

Niech $W = \{x \mid Grypa(x) = Tak\}$. Wtedy:

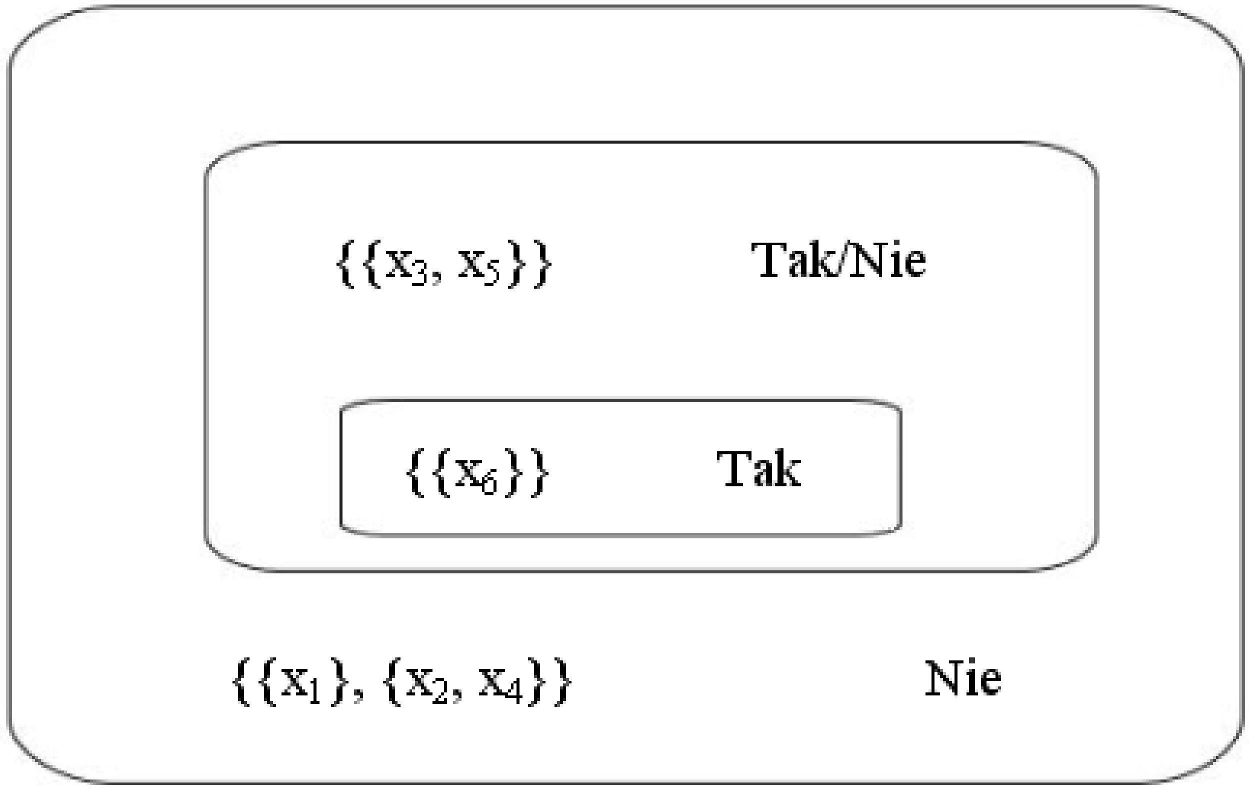
- $W = \{x_3, x_6\}$
- $\underline{A}W = \{x_6\}$
- $AW = \{x_3, x_5, x_6\}$
- $BN_A(W) = \{ x_3, x_5\}$

Przybliżanie zbioru (przykład)

Niech $W = \{x \mid Grypa(x) = Tak\}$. Wtedy:

- $W = \{x_3, x_6\}$
- $\underline{A}W = \{x_6\}$
- $AW = \{x_3, x_5, x_6\}$
- $BN_A(W) = \{ x_3, x_5\}$

Przybliżanie zbioru (przykład)



Przybliżanie zbioru (przykład 2)

	Ból głowy	**Bóle mięśni**	**Dreszcze**	**Temperatura [°C]**	***Grypa***
x_1	**Tak**	**Nie**	**Nie**	**36-37**	***Nie***
x_2	**Nie**	**Nie**	**Nie**	**36-37**	***Nie***
x_3	**Tak**	**Tak**	**Nie**	**37-38**	***Tak***
x_4	**Tak**	**Nie**	**Nie**	**36-37**	***Nie***
x_5	**Tak**	**Tak**	**Tak**	**37-38**	***Nie***
x_6	**Tak**	**Tak**	**Tak**	**38-39**	***Tak***

$U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ $[x_B] = \{\{x_1, x_4\}, \{x_2\}, \{x_3, x_5\}, \{x_6\}\}$

Niech $W = \{x \mid Grypa(x) = Tak\}$. Wtedy:

$W = \{x_3, x_6\}$

$\underline{A}W = \{x_6\}$ **(z def.** $\underline{B}X = \{x \mid [x]B \subseteq X\}$ **)**

$AW = \{x_3, x_5, X_6\}$ **(z def.** $BX = \{x \mid [x]B \cap X \neq \emptyset\}$ **)**

$BN_A(W) = \{x_3, x_5\}$ **(z def.** $BN_B(X) = BX - \underline{B}X$ **)**

Przybliżanie zbioru (przykład 2)

$U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ $[x_B] = \{\{x_1, x_4\}, \{x_2\}, \{x_3\}, \{x_5\}, \{x_6\}\}$

Niech $W = \{x \mid Grypa(x) = Tak\}$.
Wtedy:

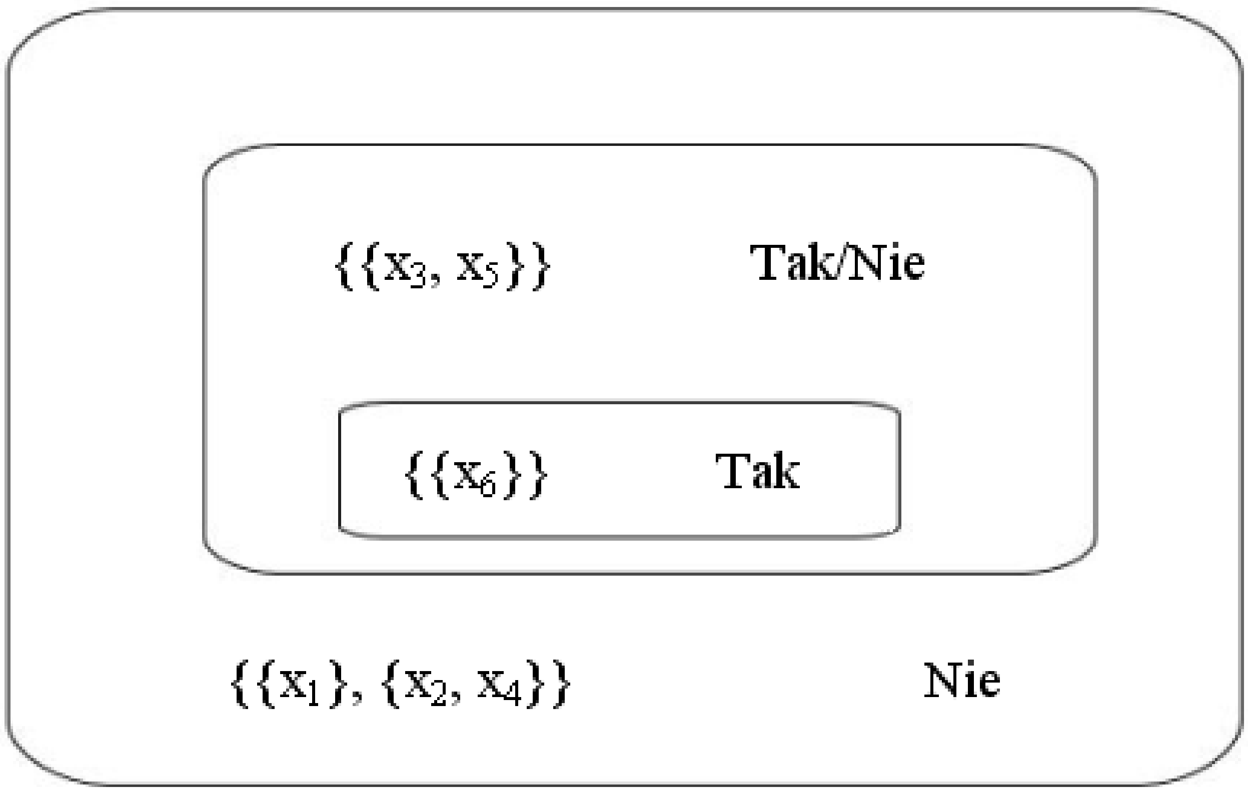
$W = \{x_3, x_6\}$

$\underline{A}W = \{x_6\}$ **(z def. $\underline{B}X = \{x \mid [x]B \subseteq X\}$)**

$\overline{A}W = \{x_3, x_5, X_6\}$ **(z def. $\overline{B}X = \{x \mid [x]B \cap X \neq \emptyset\}$)**

$BN_A(W) = \{x_3, x_5\}$ **(z def. $BNB(X) = \overline{B}X - \underline{B}X$)**

Przybliżanie zbioru (przykład 2)



Przybliżanie zbioru (przykład 2)

	Ból głowy	Bóle mięśni	Dreszcze	Temperatura [°C]	*Grypa*
x_1	Tak	Nie	Nie	36-37	*Nie*
x_2	Nie	Nie	Nie	36-37	*Nie*
x_3	Tak	Tak	Nie	37-38	*Tak*
x_4	Tak	Nie	Nie	36-37	*Nie*
x_5	Tak	Tak	Tak	37-38	*Nie*
x_6	Tak	Tak	Tak	38-39	*Tak*

$U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ $[x_B] = \{\{x_1, x_4\}, \{x_2\}, \{x_3, x_5\}, \{x_6\}\}$

Niech $W = \{x \mid Grypa(x) = Nie\}$. Wtedy:

$W = \{x_1, x_2, x_4, x_5\}$

$\underline{A}W = \{x_2\}$ **(z def.** $\underline{B}X = \{x \mid [x]B \subseteq X\}$ **)**

$\overline{A}W = \{x_1, x_2, x_3, x_4, x_5\}$ **(z def.** $\overline{B}X = \{x \mid [x]B \cap X \neq \emptyset\}$ **)**

$BN_A(W) = \{x_1, x_3, x_4, x_5\}$ **(z def.** $BN_B(X) = \overline{B}X - \underline{B}X$ **)**

Przybliżanie zbioru (przykład 2)

$U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ $[x_B] = \{\{x_1, x_4\}, \{x_2\}, \{x_3, x_5\}, \{x_6\}\}$

Niech $W = \{x \mid Grypa(x) = \textcolor{red}{Nie}\}$.
Wtedy:

$W = \{x_1, x_2, x_4, x_5\}$

$\underline{A}W = \{x_2\}$ **(z def. $\underline{B}X = \{x \mid [x]B \subseteq X\}$)**

$\overline{A}W = \{x_1, x_2, x_3, x_4, x_5\}$ **(z def. $\overline{B}X = \{x \mid [x]B \cap X \neq \emptyset\}$)**

$BN_A(W) = \{x_1, x_3, x_4, x_5\}$ **(z def. $BN_B(X) = \overline{B}X - \underline{B}X$)**

Zbiór przybliżony

Zbiór *X* nazywamy **zbiorem przybliżonym**,

jeżeli $\boldsymbol{BN_B(X) \neq \emptyset}$,

czyli jeżeli jego region graniczny jest **niepusty**.

Zbiór przybliżony

Pojęcia dzielimy na:

- **Nieostre**, które nie są pojęciami jednoznacznymi, dokładnie określonymi.
- **Ostre**, to pojęcia ściśle sklasyfikowane, przyporządkowane.

Zbiór przybliżony

Jeśli A = (U, A) jest systemem informacyjnym, B $\in$ A oraz X $\in$ U taki, że X ≠ Ø, to miarę

αB(X) = card BX/card X

będziemy nazywać **współczynnikiem dokładności** pojęcia X w systemie A, względem zbioru atrybutów B.

Zbiór przybliżony

Własności współczynnika własności:

- $0 \leq \alpha B \leq 1$,
- Jeśli **$\alpha B(X) = 1$**, to pojęcie jest **ostre** i jego własności mogą być w pełni wyrażone za pomocą zbioru atrybutów B,
- Jeśli **$\alpha B(X) = 0$** , to pojęcie X jest całkowicie **nieostre** i jego własności nie mogą być wyrażone za pomocą zbioru atrybutów B,
- Jeśli **$0 < \alpha B(X) < 1$**, to pojęcie jest również **nieostre** ale jego własności mogą być częściowo wyrażone przy pomocy atrybutów ze zbioru B.

Atrybuty zbędne i niezbędne

- **Atrybut zbędny** (zależny) – atrybut, który nie wpływa na relację nierozróżnialności. Może być pominięty.

- **Atrybut niezbędny** (niezależny) – atrybut, który wpływa na relację nierozróżnialności.

Atrybuty zbędne i niezbędne

Zbiór A (zbiór atrybutów) jest zbiorem niezależnym, jeżeli:

$\forall$ $\boldsymbol{a} \subseteq \boldsymbol{A}$, $\boldsymbol{a}$ jest niezbędny (niezależny)

Atrybuty zbędne i niezbędne

- Nie dany będzie system: $\boldsymbol{S = (U, A)}$ **i** $\boldsymbol{B \subseteq A}$
- Atrybut ***a*** jest **zbędny** (nie wpływa na relację nierozróżnialności) wtedy i tylko wtedy, gdy:

$$IND(B)=IND(B-\{a\})$$

- Atrybut ***a*** jest ***niezbędny*** (wpływa na relację nierozróżnialności) wtedy i tylko wtedy, gdy:

$$IND(B) \neq IND(B-\{a\})$$

Rdzeń

Rdzeń – jest to zbiór wszystkich niezbędnych (niezależnych) atrybutów należących do systemu decyzyjnego.

Dany system informacyjny (decyzyjny) posiada jeden rdzeń.

Rdzeń

Niech $\boldsymbol{S = (U, A)}$ **i** $\boldsymbol{B \subseteq A}$

Wtedy rdzeń zawierający wszystkie niezależne atrybuty oznaczymy jako:

CORE(B)

Redukt

Redukt – jest to minimalny zbiór niezbędnych (niezależnych) atrybutów.

Dany system informacyjny (decyzyjny) może posiadać wiele **reduktów**.

Redukty (def.)

Redukt systemu *S = (U, A)* jest definiowany jako minimalny zbiór atrybutów $\boldsymbol{B \subseteq A}$, taki że,

$$IND(B) = IND(A)$$

Zbiór wszystkich reduktów w zbiorze atrybutów A, oznaczamy:

RED(A)

Twierdzenie

Problem szukania najkrótszego reduktu jest NP-zupełny.

Redukt a Rdzeń

Zależność między reduktem, a rdzeniem jest następująca:

$$\mathbf{CORE(A) = \cap\ RED(A)}$$

Redukt zawiera w sobie rdzeń.

Atrybuty zbędne i niezbędne (przykł.1)

	Ból głowy	**Temperatura [°C]**	**Grypa**
x_1	**Tak**	**36-37**	**Nie**
x_2	**Nie**	**36-37**	**Nie**
x_3	**Tak**	**37-38**	**Tak**
x_4	**Nie**	**36-37**	**Nie**
x_5	**Tak**	**37-38**	**Nie**
x_6	**Tak**	**38-39**	**Tak**

B = {Ból głowy, Temperatura}
IND({Ból głowy}) = $\{\{x_1, x_3, x_5, x_6\}, \{x_2, x_4\}\}$
IND(Temperatura) = $\{\{x_1, x_2, x_4\}, \{x_3, x_5\}, \{x_6\}\}$
IND({Ból głowy, Temperatura}) = $\{\{x_1\}, \{x_2, x_4\}, \{x_3, x_5\}, \{x_6\}\}$

Niech
B = {***Ból głowy, Temperatura***}
oczywiście $S = (U, A)$ i $B \subseteq A$

Atrybuty zbędne i niezbędne (przykł.1)

Sprawdzamy atrybut: {Ból głowy} jest zbędny:

IND{B-{Ból głowy}} = *{{x_1, x_2 , x_4}, {x_3, x_5}, {x_6}}*
IND(B) = {{x_1}, {x_2, x_4}, {x_3, x_5}, {x_6}}

IND{B-{Ból głowy}} $\neq$ *IND(B)* $\leftarrow$ *atrybut {Ból głowy nie jest zbędny}*

Sprawdzamy, czy atrybut {Temperatura} jest zbędny:

IND{B-{Temperatura}} = *{{x_1, x_3, x_5, x_6}, {x_2 , x_4}}*
IND(B) = {{x_1}, {x_2, x_4}, {x_3, x_5}, {x_6}}

IND{B-{Temperatura}} $\neq$ *IND(B)* $\leftarrow$ *atrybut {Temperatura nie jest zbędny}*

Redukty (przykład 1)

Atrybuty B1 {Ból głowy, Temperatura} :

- Nie są zbędne (zależne)
- Nie ma innych niezbędnych atrybutów w tym systemie.

Redukty (przykład 1)

Zbiór atrybutów B1 jest minimalny tj. wszelkie podzbiory o mniejszej mocy nie spełniają warunku Reduktu

- IND({Ból głowy}) ≠ IND(B) *;*
- IND({Temperatura}) ≠ IND(B)*;*

Redukty (przykład 1)

Skoro zbiór B1 jest minimalny i:

IND({Ból głowy, Temperatura}) = IND(B)

To jest on jedynym reduktem zbioru B.

Zatem zbiór wszystkich reduktów:

RED(B) = ({Ból głowy, Temperatura})

Rdzeń (przykład 1)

Na podstawie zbioru wszystkich reduktów, można wyznaczyć rdzeń.

$$\mathbf{CORE(A)} = \cap\ \mathbf{RED(A)}$$

więc:

CORE(A) = ({Ból głowy, Temperatura})

Atrybuty zbędne i niezbędne (przykł.2)

	Ból głowy	Bóle mięśni	Dreszcze	Temperatura [°C]	*Grypa*
x_1	Tak	Nie	Nie	36-37	*Nie*
x_2	Nie	Nie	Nie	36-37	*Nie*
x_3	Tak	Tak	Nie	37-38	*Tak*
x_4	Tak	Nie	Nie	36-37	*Nie*
x_5	Tak	Tak	Tak	37-38	*Nie*
x_6	Tak	Tak	Tak	38-39	*Tak*

IND({Ból głowy}) = $\{\{x_1, x_3, x_4, x_5, x_6\}, \{x_2\}\}$
IND({Bóle mięśni}) = $\{\{x_3, x_5, x_6\}, \{x_1, x_2, x_4\}\}$
IND({Dreszcze}) = $\{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6\}\}$
IND(Temperatura) = $\{\{x_1, x_2, x_4\}, \{x_3, x_5\}, \{x_6\}\}$
IND({Ból głowy, Bóle mięśni, Dreszcze, Temperatura}) = $\{\{x_1, x_4\}, \{x_2\}, \{x_3\}, \{x_5\}, \{x_6\}\}$

Niech
B = {*Ból głowy, Bóle mięśni, Dreszcze, Temperatura*}
oczywiście $\boldsymbol{S = (U, A)}$ **i** $\boldsymbol{B \subseteq A}$

Atrybuty zbędne i niezbędne (przykł.2)

Sprawdzamy atrybut: {Ból głowy} jest zbędny:

IND{B-{Ból głowy}} = IND({Bóle mięśni, Dreszcze, Temperatura})

IND{B-{Ból głowy}} = $\{\{x_1, x_2, x_4\}, \{x_3, x_5\}, \{x_6\}\}$

IND(B) = $\{\{x_1, x_4\}, \{x_2\}, \{x_3\}, \{x_5\}, \{x_6\}\}$

IND{B-{Ból głowy}} $\neq$ *IND(B)* ← *atrybut {Ból głowy nie jest zbędny}*

Sprawdzamy, czy atrybut {Bóle mięśni} jest zbędny:

IND{B-{Bóle mięśni}} = IND({Bóle głowy, Dreszcze, Temperatura})

IND{B-{Bóle mięśni}} = $\{\{x_1, x_4\}, \{x_2\}, \{x_3\}, \{x_5\}, \{x_6\}\}$

IND(B) = $\{\{x_1, x_4\}, \{x_2\}, \{x_3\}, \{x_5\}, \{x_6\}\}$

IND{B-{Bóle mięśni}} = *IND(B)* ← *atrybut {Bóle mięśni jest zbędny}*

Atrybuty zbędne i niezbędne (przykł.2)

Sprawdzamy atrybut: czy {Dreszcze} jest zbędny:

IND{B-{Dreszcze}} = IND({Bóle głowy, Bóle mięśni, Temperatura})

IND{B-{Dreszcze}} = *{{x_1, x_4}, {x_2}, {x_3, x_4}, {x_6}}*

IND(B) = {{x_1, x_4}, {x_2}, {x_3}, {x_5}, {x_6}}

IND{B-{Dreszcze}} ≠ *IND(B)* ← *atrybut {*Dreszcze *nie jest zbędny}*

Sprawdzamy, czy atrybut {Temperatura} jest zbędny:

IND{B-{Temperatura}} = IND({Bóle głowy, Bóle mięśni, Temperatura})

IND{B-{Temperatura}} = *{{x_1, x_4}, {x_2}, {x_3, x_5, x_6}}*

IND(B) = {{x_1, x_4}, {x_2}, {x_3}, {x_5}, {x_6}}

IND{B-{Temperatura}} ≠ *IND(B)* ← *atrybut {Temperatura nie jest zbędny}*

Redukty (przykład 2)

Atrybuty B1{Ból głowy, Dreszcze, Temperatura} :

- Nie są zbędne (zależne)
- Nie ma innych niezbędnych atrybutów w tym systemie.

Atrybut {Bóle mięśni} :

- Jest zbędny (Nie wpływa na rozróżnialności obiektów Uniwersum)

Redukty (przykład 2)

Zbiór atrybutów B1 jest minimalny tj. wszelkie podzbiory o mniejszej mocy nie spełniają warunku Reduktu

IND({Ból głowy}) ≠ IND(B) *;*

IND({Dreszcze}) ≠ IND(B)*;*

IND({Temperatura}) ≠ IND(B)*;*

IND({Ból głowy, Dreszcze}) ≠ IND(B);

IND({Ból głowy, Temperatura}) ≠ IND(B)*;*

IND({Dreszcze, Temperatura}) ≠ IND(B);

Redukt (przykład 2)

Skoro zbiór B1 jest minimalny i:

IND({Ból głowy, Dreszcze, Temperatura}) = IND(B)

To jest on jedynym reduktem zbioru B.

Zatem zbiór wszystkich reduktów:

RED(B) = ({Ból głowy, Dreszcze, Temperatura})

Rdzeń (przykład 2)

Na podstawie zbioru wszystkich reduktów, można wyznaczyć rdzeń.

$$\mathbf{CORE(A) = \cap\ RED(A)}$$

więc:

CORE(A) = ({Ból głowy, Dreszcze, Temperatura})

Redukty (przykład 2.1)

	a	b	c
x_1	1	3	3
x_2	2	2	2
x_3	3	1	3

Niech S = (U, A) i B $\subseteq$ ***A***

$IND(B) = \{\{x_1\},\{x_2\},\{x_3\}\}$

$IND(B-\{a\}) = \{\{x_1,x_2\},\{x_3\}\} \neq IND(B)$ ← atrybut niezbędny

$IND(B-\{b\}) = \{\{x_1\},\{x_2,x_3\}\} \neq IND(B)$ ← atrybut niezbędny

$IND(B-\{c\}) = \{\{x_1\},\{x_2\},\{x_3\}\} = IND(B)$ ← *atrybut zbędny*

Redukty (przykład 2.1)

	a	b	c
x_1	1	3	3
x_2	2	2	2
x_3	3	1	3

Niech S = (U, A)* i *B* $\subseteq$ *A

***IND(B)* = $\{\{x_1\},\{x_2\},\{x_3\}\}$**

IND({a}) = $\{\{x_1\},\{x_2\},\{x_3\}\}$ = IND(B) ← Redukt

IND({b}) = $\{\{x_1,x_2\},\{x_3\}\}$ = IND(B) ← Redukt

IND({a,b}) = $\{\{x_1\},\{x_2\},\{x_3\}\}$ ≠ IND(B)

RED(B) = {{a},{b}} ← Zbiór wszystkich reduktów

Rdzeń (przykład 2.1)

	a	b	c
X_1	1	3	3
X_2	2	2	2
X_3	3	1	3

Niech S = (U, A) i B $\subseteq$ A

Skoro RED(B) = {{a},{b}} , to

CORE(B) = $\bigcap$ RED(B) = $\bigcap$ {{a},{b}} = $\varnothing$

Macierz rozróżnialności

Macierz rozróżnialności (ang. discernability matrix) *M* jest macierzą o wielkości *n x n*, o elementach m_{ij}, takich że:

$$m_{ij} = \{a \in A: a(x_i) \neq a(x_j)\} \text{ dla } i, j = 1,\dots,n$$

Macierz rozróżnialności ma za zadanie podsumować relacje nierozróżnialności zachodzące w systemie **S.**

Macierz rozróżnialności (przykład)

b – ból głowy, t - temperatura

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-					
x_2	b	-				
x_3	t	b, t	-			
x_4	b	$\varnothing$	b, t	-		
x_5	t	b, t	$\varnothing$	b, t	-	
x_6	t	b, t	t	b, t	t	-

Macierz rozróżnialności (przykład 2)

g – ból głowy, m- ból mięśni, d- dreszcze t - temperatura

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-					
x_2	g	-				
x_3	m,t	g,m,t	-			
x_4	$\varnothing$	g	m,t	-		
x_5	m,d,t	g,m,d,t	d	m,d,t	-	
x_6	m,d,t	g,m,d,t	d,t	m,d,t	t	-

Macierz rozróżnialności

Na podstawie Macierzy Rozróżnialności, możemy łatwo wyznaczyć atrybuty niezbędne.

Są to elementy występujące w macierzy jednostkowo (pojedyńczo)

Macierz rozróżnialności (przykład 1,2)

W przypadku rozważanej macierzy (przykład 1) będą to atrybuty:

b , **t**

W przypadku rozważanej macierzy (przykład 2) będą to atrybuty:

g , **d** **,** **t**

Macierz rozróżnialności

atrybuty niezbędne(przykład 1,2)

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-					
x_2	b	-				
x_3	t	b, t	-			
x_4	b	$\varnothing$	b, t	-		
x_5	t	b, t	$\varnothing$	b, t	-	
x_6	t	b, t	t	b, t	t	-

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-					
x_2	g	-				
x_3	m, t	g, m, t	-			
x_4	$\varnothing$	g	m, t	-		
x_5	m, d,t	g, m, d,t	d	m, d,t	-	
x_6	m, d,t	g, m, d,t	d,t	m, d,t	t	-

Macierz rozróżnialności

Na podstawie Macierzy Rozróżnialności, możemy łatwo wyznaczyć redukty.

Są to elementy mające niepustą część wspólną z każdym niepustym elementem należącym do macierzy.

Macierz rozróżnialności (przykład 1,2)

W przypadku rozważanej macierzy (przykład 1) będą to atrybuty:

b , **t**

W przypadku rozważanej macierzy (przykład 2) redukt stanowiły będą atrybuty:

g , **d , t**

Macierz rozróżnialności
redukty(przykład 1,2)

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-					
x_2	b	-				
x_3	t	b, t	-			
x_4	b	∅	b, t	-		
x_5	t	b, t	∅	b, t	-	
x_6	t	b, t	t	b, t	t	-

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-					
x_2	g	-				
x_3	m,t	g,m,t	-			
x_4	∅	g	m,t	-		
x_5	m,d,t	g,m,d,t	d	m,d,t	-	
x_6	m,d,t	g,m,d,t	d,t	m,d,t	t	-

Funkcja rozróżnialności

Funkcja rozróżnialności jest funkcją Boole'a o m zmiennych Boole'a $a_1^*,\ldots,a_m^*$ (nawiązujących do atrybutów $a_1,\ldots,a_m$, gdzie $c_{i,j}^*=\{a^*:a\in c_{ij}\}$.

$f_S(a_1^*,\ldots,a_m^*) = \wedge\{\vee c_{i,j}^*: 1 \leq j \leq j \leq n, c_{i,j} \neq \emptyset\}$

Funkcja rozróżnialności

Funkcja rozróżnialności jest prostą koniunkcją wszystkich elementów macierzy M.

Koniunkcja ta może zostać uproszczona przy pomocy algebry Boole'a, a jej wynikiem są możliwe **redukty** systemu.

Funkcja rozróżnialności (przykład 1)

Korzystamy z macierzy rozróżnialności z poprzedniego przykładu

$$
\begin{aligned}
f_S(b, t) = \quad & (b) \wedge (t) \wedge (b) \wedge (t) \wedge (t) \wedge \\
& (b \vee t) \wedge (b \vee t) \wedge (b \vee t) \wedge \\
& (b \vee t) \wedge (t) \wedge \\
& (b \vee t) \wedge (b \vee t) \wedge \\
& (t) = (b) \wedge (t)
\end{aligned}
$$

Czyli w tym przypadku **nie możemy** zmniejszyć rozmiaru danych przez usunięcie atrybutu. Oba atrybuty (ból głowy, temperatura) muszą zostać.

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-					
x_2	b	-				
x_3	t	b, t	-			
x_4	b	∅	b, t	-		
x_5	t	b, t	∅	b, t	-	
x_6	t	b, t	t	b, t	t	-

Funkcja rozróżnialności(przykład 2)

Korzystamy z poniższej macierzy rozróżnialności, opisujący przykład nr 2:

$$f_S(g, m, d, t) = g \wedge (m \vee t) \wedge (g \vee m \vee t) \wedge g \wedge (m \vee t) \wedge (m \vee d \vee t)$$

$$\wedge (g \vee m \vee d \vee t) \wedge d \wedge (m \vee d \vee t) \wedge (m \vee d \vee t)$$

$$\wedge (g \vee m \vee d \vee t) \wedge (d \vee t) \wedge (m \vee d \vee t) \wedge t$$

$$= g \wedge d \wedge t$$

W tym przypadku możemy usunąć atrybut m (ból mięśni)

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-					
x_2	g	-				
x_3	m,t	g,m,t	-			
x_4	$\varnothing$	g	m,t	-		
x_5	m,d,t	g,m,d,t	d	m,d,t	-	
x_6	m,d,t	g,m,d,t	d,t	m,d,t	t	-

Funkcja rozróżnialności (przykład 3)

Tabela przedstawia system decyzyjny z kandydatami, ich atrybutami (wykształcenie, doświadczenie, znajomość języka angielskiego, referencje) oraz decyzje czy danego kandydata należy przyjąć, czy odrzucić.

	Wykształcenie	Doświadczenie	Angielski	Referencje	Decyzja
x_1	Wyższe	Średnie	Tak	Znakomite	Tak
x_2	Wyższe	Małe	Tak	Neutralne	Nie
x_3	Podstawowe	Małe	Tak	Dobre	Nie
x_4	Średnie	Duże	Tak	Neutralne	Tak
x_5	Średnie	Średnie	Tak	Neutralne	Nie
x_6	Średnie	Duże	Tak	Znakomite	Tak
x_7	Wyższe	Duże	Nie	Dobre	Tak
x_8	Podstawowe	Małe	Nie	Znakomite	Nie

Funkcja rozróżnialności (przykład 2.)

Kilka przykładowych relacji nierozróżnialności:

- *IND({Wykształcenie}) =*

  $\{\{x_1, x_2, x_7\}, \{x_3, x_8\}, \{x_4, x_5, x_6\}\}$
- *IND({Wykształcenie, Angielski})=*

  $\{\{x_1, x_2\}, \{x_3\}, \{x_4, x_5, x_6\}, \{x_7\}, \{x_8\}\}$
- *IND({Angielski, Referencje}) =*

  $\{\{x_1, x_6\}, \{x_2, x_4, x_5\}, \{x_3\}, \{x_7\}, \{x_8\}\}$
- *IND({W, D, A, R})=*

  $\{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}\}$

Funkcja rozróżnialności (przykład 2.)

Wyznaczmy **dolne** oraz **górne** przybliżenie:

Niech $W = \{x \mid Decyzja(x) = Tak\}$. Wtedy:

- $W = \{x_1, x_4, x_6, x_7\}$
- $\underline{A}W = \{x_1, x_4, x_6, x_7\}$
- $AW = \{x_1, x_4, x_6, x_7\}$

Funkcja rozróżnialności (przykład 3)

Macierz rozróżnialności:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	-							
x_2	d,r	-						
x_3	w,d,r	w,r	-					
x_4	w,d,r	w,d	w,d,r	-				
x_5	w,r	w,d	w,d,r	d	-			
x_6	w,d	w,d,r	w,d,r	r	d,r	-		
x_7	d,a,r	d,a,r	w,d,a	w,a,r	w,d,a,r	w,a,r	-	
x_8	w,d,a	w,a,r	a,r	w,d,a,r	w,d,a,r	w,d,a	w,d,r	-

Funkcja rozróżnialności (przykład 2.)

Zbudujmy funkcję rozróżnialności:

$$
\begin{aligned}
f_S(w, d, a, r) = & (d \vee r) \wedge (w \vee d \vee r) \wedge (w \vee d \vee r) \wedge (w \vee r) \\
& \wedge (w \vee d) \wedge (d \vee a \vee r) \wedge (w \vee d \vee a) \\
& \wedge (w \vee r) \wedge (w \vee d) \wedge (w \vee d) \wedge (w \vee d \vee r) \\
& \wedge (d \vee a \vee r) \wedge (w \vee a \vee r) \wedge (w \vee d \vee r) \\
& \wedge (w \vee d \vee r) \wedge (w \vee d \vee r) \wedge (w \vee d \vee a) \\
& \wedge (a \vee r) \wedge (d) \wedge (r) \wedge (w \vee a \vee r) \\
& \wedge (w \vee d \vee a \vee r) \wedge (d \vee r) \wedge (w \vee d \vee a \vee r) \\
& \wedge (w \vee d \vee a \vee r) \wedge (w \vee a \vee r) \wedge (w \vee d \vee a) \\
& \wedge (w \vee d \vee r)
\end{aligned}
$$

Funkcja rozróżnialności (przykład 3)

Po uproszczeniu funkcja $\boldsymbol{f_S(w, d, a, r) = d \wedge r}$. Czyli 'Doświadczenie' i 'Referencje' są **reduktami**. Możemy to sprawdzić przez porównanie:

$IND(\{W, D, A, R\})$ z $IND(\{D, R\})$.

$IND(\{W, D, A, R\})$

$= \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{ x_7\},\{ x_8\}\}$

$IND(\{D,R\})$

$= \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{ x_7\},\{ x_8\}\}$

Czyli $\boldsymbol{IND(\{W, D, A, R\}) = IND(\{D, R\})}$.

Niespójności

- W każdym układzie decyzyjnym, mogą pojawić się elementy Uniwersum o tych samych wartościach atrybutów, lecz o innej wartości funkcji decyzyjnej.

	Angielski	Java	Doświadczenie	Rekrutacja
x_1	Nie	Tak	Niskie	Tak
x_2	Nie	Tak	Niskie	Nie

Niespójności - przykład

	Angielski	Java	Doświadczenie	Rekrutacja
$\mathbf{x}_1$	Nie	Tak	Niskie	Tak
$\mathbf{x}_2$	Nie	Tak	Niskie	Nie
$\mathbf{x}_3$	Tak	Nie	Średnie	Nie
$\mathbf{x}_4$	Tak	Tak	Średnie	Tak
$\mathbf{x}_5$	Tak	Tak	Wysokie	Tak

Niespójności - usuwanie

- Istnieją różne metody usuwania niespójności, np.:
  - Metoda Jakościowa
  - Metoda Nowego Podziału Systemu Informacyjnego

Niespójności - usuwanie

- Metoda jakościowa:

  Z systemu o niespójnej wiedzy, należy usunąć ten obiekt, dla którego dokładność górnego bądź dolnego przybliżenia jest mniejsza.

Niespójności - przykład

	Angielski	Java	Doświadczenie	Rekrutacja
x_1	Nie	Tak	Niskie	Tak
x_2	Nie	Tak	Niskie	Nie
x_3	Tak	Nie	Średnie	Nie
X_4	Tak	Tak	Średnie	Tak
x_5	Tak	Tak	Wysokie	Tak

Niech S = (U, A)* i *B* $\subseteq$ *A

W1 = {x | Rekrutacja(x) = Tak}.

$W1 = \{x_1, x_4, x_5\}$

$\underline{A}W1 = \{x_1, x_4, x_5\}$

$AW1 = \{x_1, x_2, x_3, x_4, x_5\}$

W2 = {x | Rekrutacja(x) = Nie}.

$W2 = \{x_2, x_3\}$

$\underline{A}W2 = \{x_2, x_4, x_5\}$

$AW2 = \{x_1, x_2, x_3, x_4, x_5\}$

Niespójności - przykład

Niech S = (U, A)* i *B* $\subseteq$ *A

$$\boldsymbol{U} = \{x_1, x_2, x_3, x_4, x_5\}$$

$$\boldsymbol{IND(B)} = \{\{x_1, x_2\}, \{x_3\}, \{x_4\}, \{x_5\}\}$$

$W1 = \{x \mid Rekrutacja(x) = Tak\}$.

$W1 = \{x_1, x_4, x_5\}$

$\underline{A}W1 = \{x_4, x_5\}$

$AW1 = \{x_1, x_2, x_4, x_5\}$

$W2 = \{x \mid Rekrutacja(x) = Nie\}$.

$W2 = \{x_2, x_3\}$

$\underline{A}W2 = \{x_3\}$

$AW2 = \{x_1, x_2, x_3\}$

Niespójności - przykład

Dokładność przybliżenia górnego: $P = \frac{|\underline{B}X|}{|U|}$

$P = \frac{|AW1|}{5} = \frac{4}{5}$ $\qquad$ $P = \frac{|AW2|}{5} = \frac{3}{5}$

Dokładność przybliżenia dolnego: $P = \frac{|BX|}{|U|}$

$P = \frac{|\underline{A}W1|}{5} = \frac{2}{5}$ $\qquad$ $P = \frac{|\underline{A}W2|}{5} = \frac{1}{5}$

Niespójności - przykład

Przybliżenia są mniejsze dla zbioru: W2 = {x | Rekrutacja(x) = Nie}

Usuwamy z niego obiekt $\mathbf{x_2}$, który wraz z obiektem $\mathbf{x_1}$ stanowił niespójność.

Nowa tabela jest teraz spójna.

	Angielski	Java	Doświadczenie	Rekrutacja
$\mathbf{x_1}$	Nie	Tak	Niskie	Tak
$\mathbf{x_3}$	Tak	Nie	Średnie	Nie
$\mathbf{X_4}$	Tak	Tak	Średnie	Tak
$\mathbf{x_5}$	Tak	Tak	Wysokie	Tak

Generowanie reguł

- Reguła decyzyjna – jest to obiekt należący do tablicy decyzyjnej, zapisany w postaci:

Jeśli {warunek bądź warunki} to {decyzja}

Warunek reguły decyzyjnej

Decyzja reguły decyzyjnej

Generowanie reguł (przykład 1)

	Ból głowy	Temperatura [°C]	Grypa
x_1	Tak	36-37	Nie
x_2	Nie	36-37	Nie
x_3	Tak	37-38	Tak
x_4	Nie	36-37	Nie
x_5	Tak	37-38	Nie
x_6	Tak	38-39	Tak

Reguły decyzyjne:

IF {Ból głowy == Tak} AND {Temperatura == 36-37} **THEN** Grypa = Nie
IF {Ból głowy == Nie} AND {Temperatura == 36-37} **THEN** Grypa = Nie
IF {Ból głowy == Tak} AND {Temperatura == 37-38} **THEN** Grypa = Tak
IF {Ból głowy == Nie} AND {Temperatura == 36-37} **THEN** Grypa = Nie
IF {Ból głowy == Tak} AND {Temperatura == 37-38} **THEN** Grypa = Nie
IF {Ból głowy == Tak} AND {Temperatura == 38-39} **THEN** Grypa = Tak

Generowanie reguł (przykład 2)

	Ból głowy	Bóle mięśni	Dreszcze	Temperatura [°C]	*Grypa*
x_1	Tak	Nie	Nie	36-37	*Nie*
x_2	Nie	Nie	Nie	36-37	*Nie*
x_3	Tak	Tak	Nie	37-38	*Tak*
x_4	Tak	Nie	Nie	36-37	*Nie*
x_5	Tak	Tak	Tak	37-38	*Nie*
x_6	Tak	Tak	Tak	38-39	*Tak*

Niech: Ból głowy: BG, Bóle mięśni: BM, Dreszcze: D, Temperatura: T

Reguły decyzyjne:

IF {BG == Tak} AND {BM == Nie} AND {D == Nie} AND {T == 36-37} **THEN** Grypa = Nie
IF {BG == Nie} AND {BM == Nie} AND {D == Nie} AND {T == 36-37} **THEN** Grypa = Nie
IF {BG == Tak} AND {BM == Tak} AND {D == Nie} AND {T == 36-37} **THEN** Grypa = Tak
IF {BG == Nie} AND {BM == Nie} AND {D == Nie} AND {T == 36-37} **THEN** Grypa = Nie
IF {BG == Tak} AND {BM == Tak} AND {D == Tak} AND {T == 36-37} **THEN** Grypa = Nie
IF {BG == Tak} AND {BM == Tak} AND {D == Tak} AND {T == 36-37} **THEN** Grypa = Tak

Generowanie reguł minimalnych

Możemy wygenerować również reguły minimalne.

Musimy w tym celu wygenerować macierz rozróżnialności dla wierszy o różnych wartościach decyzji.

Generowanie reguł minimalnych

Definiujemy macierz:

$$m_{ij} = \{a \in A: a(x_i) \neq a(x_j)\} \text{ dla } i, j = 1,\ldots,n$$

$$M_{ij} = \begin{cases} \{a \in A: a(x_i) \neq a(x_j)\} \text{ dla } i,j=1,\ldots,n; & \text{jeśli } d(x_i) \neq d(x_j) \\ \lambda & \text{; jeśli } d(x_i) = d(x_j) \end{cases}$$

Gdzie, λ – wartość wstawiana w przypadku wierszy o takich samych decyzjach d(wiersz)

Generowanie reguł minimalnych (przykład 1)

	$\mathbf{x_1}$	$\mathbf{x_2}$	$\mathbf{x_3}$	$\mathbf{x_4}$	$\mathbf{x_5}$	$\mathbf{x_6}$
$\mathbf{x_1}$	-	λ	t	λ	λ	t
$\mathbf{x_2}$	λ	-	b,t	λ	λ	b,t
$\mathbf{x_3}$	t	b,t	-	b,t	$\emptyset$	λ
$\mathbf{x_4}$	λ	λ	b,t	-	λ	b,t
$\mathbf{x_5}$	λ	λ	$\emptyset$	λ	-	t
$\mathbf{x_6}$	t	b,t	λ	b,t	t	-

Funkcje rozróżnialności dla wierszy:

F(A,{Nie}, $\mathbf{X_1}$) = $t \wedge t = t$
F(A,{Nie}, $\mathbf{X_2}$) = $(b \vee t) \wedge (b \vee t) = b \vee t$
F(A,{Tak}, $\mathbf{X_3}$) = $t \wedge (b \vee t) \wedge (b \vee t) = t$
F(A,{Nie}, $\mathbf{X_4}$) = $(b \vee t) \wedge (b \vee t) = b \vee t$
F(A,{Nie}, $\mathbf{X_5}$) = t
F(A,{Tak}, $\mathbf{X_6}$) = $t \wedge (b \vee t) \wedge (b \vee t) \wedge t = t$

Generowanie reguł minimalnych (przykład 1)

Reguły minimalne

Dla $F(A,\{Nie\}, \mathbf{X_1}) = t$

IF Temperatura == 36-37 THEN Grypa = Nie

Dla $F(A,\{Nie\}, \mathbf{X_2}) = b \vee t$

IF Ból głowy == Nie THEN Grypa = Nie

IF Temperatura == 36-37 THEN Grypa = Nie

Dla $F(A,\{Tak\}, \mathbf{X_3}) = t$

IF Temperatura == 37-38 THEN Grypa = Tak

Generowanie reguł minimalnych (przykład 1)

Reguły minimalne

Dla F(A,{Nie}, $\mathbf{X_4}$) = b $\vee$ *t*

IF Ból głowy == Nie THEN Grypa = Nie
IF Temperatura == 36-37 THEN Grypa = Nie

Dla F(A,{Nie}, $\mathbf{X_5}$) = *t*

IF Temperatura == 37-38 THEN Grypa = Nie

Dla F(A,{Tak}, $\mathbf{X_6}$) = *t*

IF Temperatura == 38-39 THEN Grypa = Tak

Generowanie reguł (przykład 1)

Dzięki pojedyńczym regułom, można wyznaczyć ogólną regułę dla decyzji Grypa = Nie

IF $(t == 36\text{-}37)$
$\vee (b == Nie) \vee (t == 36\text{-}37)$
$\vee (b == Nie) \vee (t == 36\text{-}37)$
$\vee (t == 37\text{-}38)$
THEN Grypa = Nie

Generowanie reguł (przykład 1)

po uproszczeniu:

IF $(t == 36\text{-}37) \vee (b == Nie) \vee (t == 37\text{-}38)$
THEN Grypa = Nie

ostatecznie:

IF (Temperatura == 36-37) $\vee$ (Ból głowy == Nie) $\vee$ (Temperatura == 37-38)
THEN Grypa = Nie

Generowanie reguł (przykład 1)

Dzięki pojedyńczym regułom, można wyznaczyć ogólną regułę dla decycji Grypa = Nie

IF $(t == 37\text{-}38) \vee (t == 38\text{-}39)$
THEN Grypa = Tak

ostatecznie:

IF (Temperatura == 37-38) OR (Temperatura == 38-39)
THEN Grypa = Tak

Generowanie reguł minimalnych

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-	λ	m,t	λ	λ	m,d,t
x_2	λ	-	g,m,t	λ	λ	g,m,d,t
x_3	m,t	g,m,t	-	m,t	d	λ
x_4	λ	λ	m,t	-	λ	m,d,t
x_5	λ	λ	d	λ	-	t
x_6	m,d,t	g,m,d,t	λ	m,d,t	t	-

Przykład 2

Funkcje rozróżnialności dla wierszy:

F(A,{Nie}, $\mathbf{X_1}$) = $(m \vee t) \wedge (m \vee d \vee t) = m \vee t$

F(A,{Nie}, $\mathbf{X_2}$) = $(g \vee m \vee t) \wedge (g \vee m \vee d \vee t) = g \vee m \vee t$

F(A,{Tak}, $\mathbf{X_3}$) = $(m \vee t) \wedge (g \vee m \vee t) \wedge (m \vee t) \wedge d = m \vee t$

F(A,{Nie}, $\mathbf{X_4}$) = $(m \vee t) \wedge (m \vee d \vee t) = m \vee t$

F(A,{Nie}, $\mathbf{X_5}$) = $d \wedge t$

F(A,{Tak}, $\mathbf{X_6}$) = $(m \vee d \vee t) \wedge (g \vee m \vee d \vee t) \wedge (m \vee d \vee t) \wedge t =$
$m \vee d \vee t$

Generowanie reguł minimalnych (przykład 2)

Reguły minimalne

Dla F(A,{Nie}, $\mathbf{X_1}$) = $m \vee t$

IF Bóle mięśni == Nie THEN Grypa = Nie
IF Temperatura == 36-37 THEN Grypa = Nie

Dla F(A,{Nie}, $\mathbf{X_2}$) = $g \vee m \vee t$

IF Ból głowy == Nie THEN Grypa = Nie
IF Bóle mięśni == Nie THEN Grypa = Nie
IF Temperatura == 36-37 THEN Grypa = Nie

Dla F(A,{Tak}, $\mathbf{X_3}$) = $m \vee t$

IF Bóle mięśni == Tak THEN Grypa = Tak

IF Temperatura == 37-38 THEN Grypa = Tak

Generowanie reguł minimalnych (przykład 2)

Reguły minimalne

Dla F(A,{Nie}, $\mathbf{X}_4$) = $m \vee t$

IF Bóle mięśni == Nie THEN Grypa = Nie

IF Temperatura == 36-37 THEN Grypa = Nie

Dla F(A,{Nie}, $\mathbf{X}_5$) = $d \wedge t$

IF Dreszcze == Tak AND Temperatura == 37-38 THEN Grypa = Nie

Dla F(A,{Tak}, $\mathbf{X}_6$) = $m \vee d \vee t$

IF Bóle mięśni == Tak THEN Grypa = Tak

IF Dreszcze == Tak THEN Grypa = Tak

IF Temperatura == 38-39THEN Grypa = Tak

Generowanie reguł (przykład 2)

Dzięki pojedyńczym regułom, można wyznaczyć ogólną regułę dla decycji Grypa = Nie

IF $(m == Nie) \lor (t == 36\text{-}37)$
$\lor (g == Nie) \lor (m == Nie) \lor (t == 36\text{-}37)$
$\lor (m == Nie) \lor (t == 36\text{-}37)$
$\lor ((d == Tak) \land (t == 38\text{-}39))$
THEN Grypa = Nie

Generowanie reguł (przykład 2)

po uproszczeniu:

IF (*m* ==*Nie*) $\vee$ (*t* == *36-37*) $\vee$ (*g* == *Nie*)$\vee$ [(*d* == *Tak*)$\wedge$ (*t* == *38-39*)]
THEN Grypa = Nie

Ostatecznie:

IF (*Bóle mięśni* ==*Nie*) $\vee$ (*Temperatura* == *36-37*)
$\vee$ (*Bóle głowy* == *Nie*)
$\vee$ **[**(*Dreszcze* == *Tak*)$\wedge$ (*Temperatura* == *38-39*)**]**
THEN Grypa = Nie

Generowanie reguł (przykład 2)

Dzięki pojedyńczym regułom, można wyznaczyć ogólną regułę dla decyzji Grypa = Tak

IF $(m == Tak) \vee (t == 37\text{-}38)$
$\vee (m == Tak) \vee (d == Tak) \vee (t == 38\text{-}39)$

THEN Grypa = Tak

Generowanie reguł minimalnych (przykład 2)

po uproszczeniu:

IF (*m ==Tak*) $\vee$ *(t == 37-38)* $\vee$ (*d == Tak*) $\vee$ *(t == 38-39)*

THEN Grypa = Tak

ostatecznie:

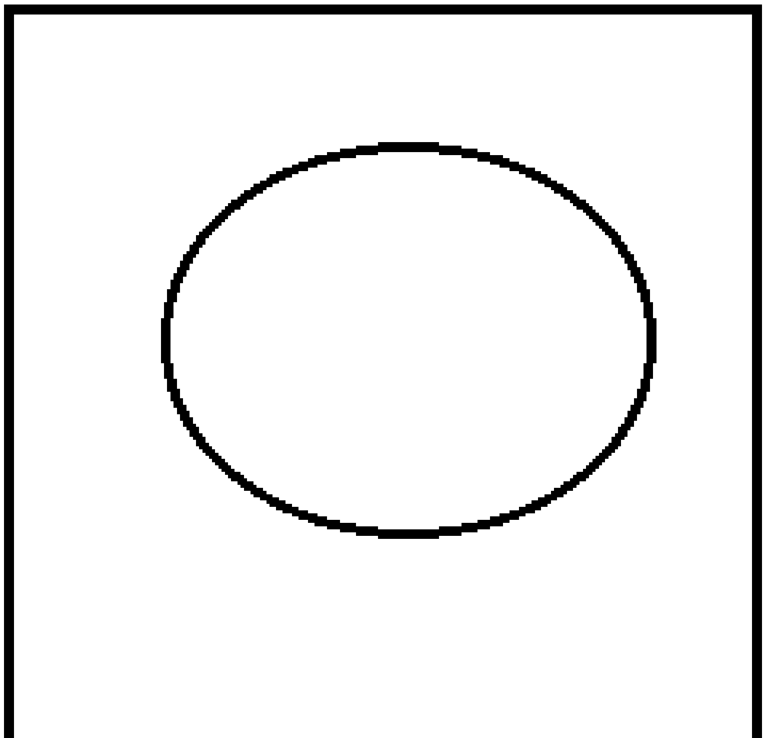
IF (Bóle mięśni == Tak) $\vee$ *(Temperatura = 37-38)* $\vee$ *(Dreszcze = Tak)*

$\vee$ *(Temperatura = 38-39)*

THEN Grypa = Tak

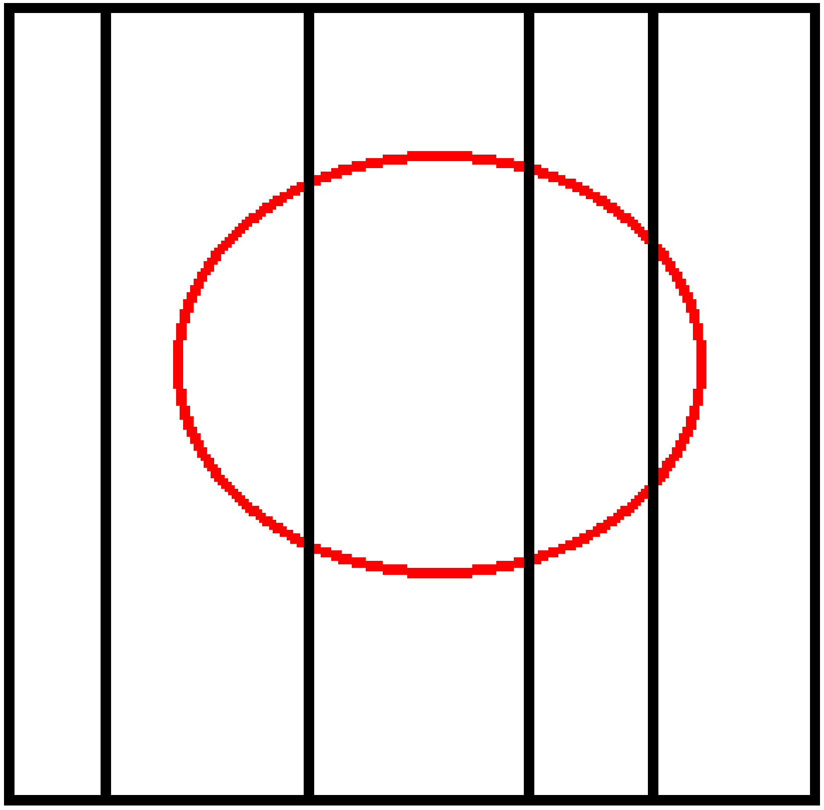
Podsumowanie

Zbiór do przybliżenia, brak użytych atrybutów



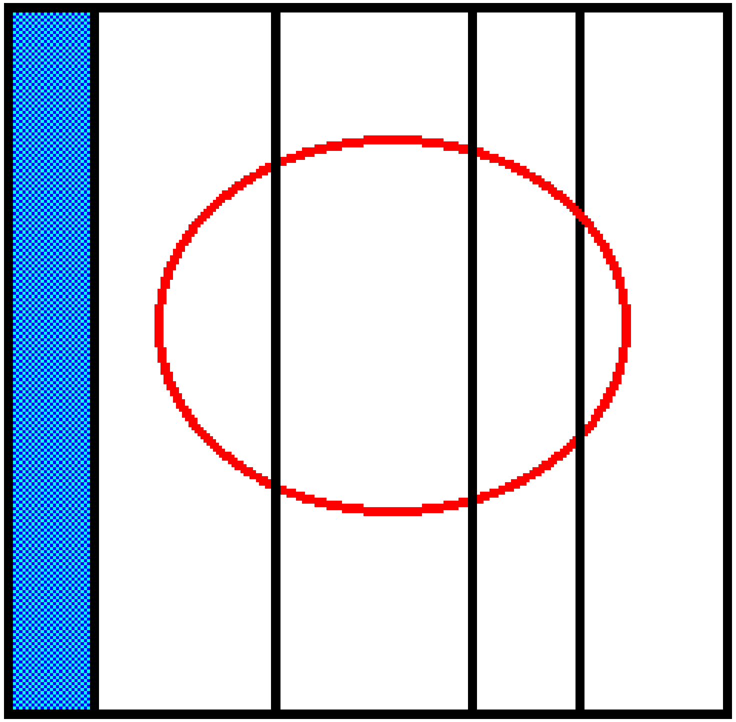
Podsumowanie

Podział z pierwszym atrybutem



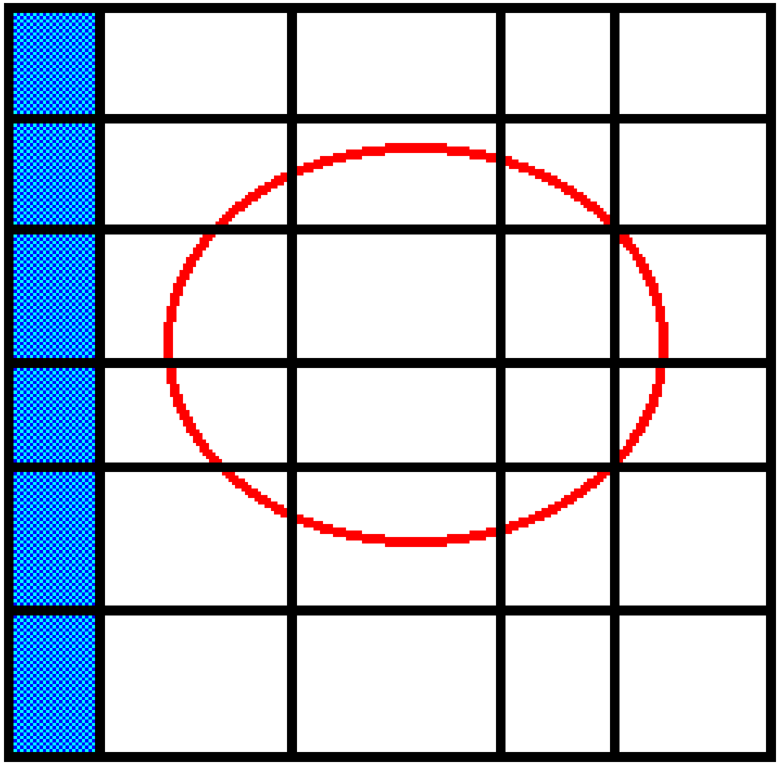
Podsumowanie

Obszar negatywny



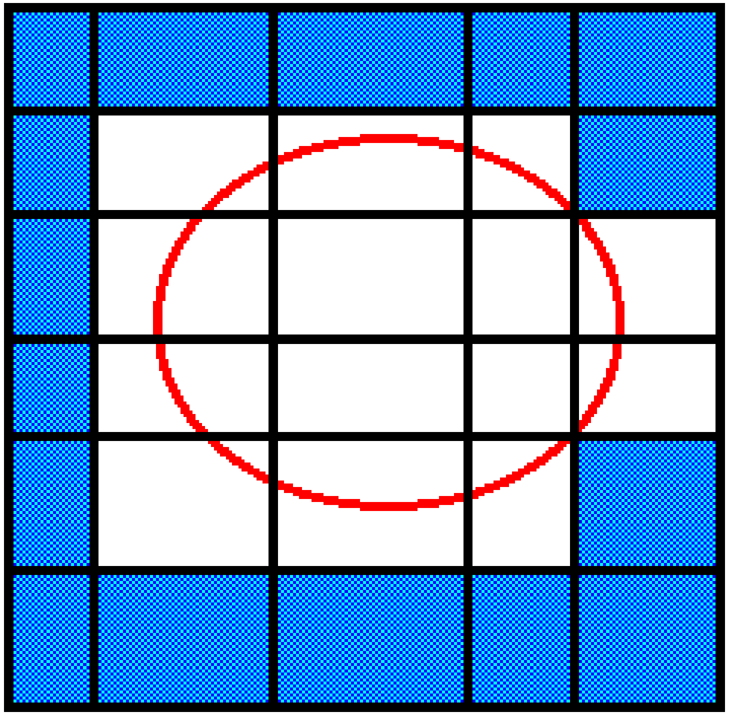
Podsumowanie

Podział z dwoma atrybutami



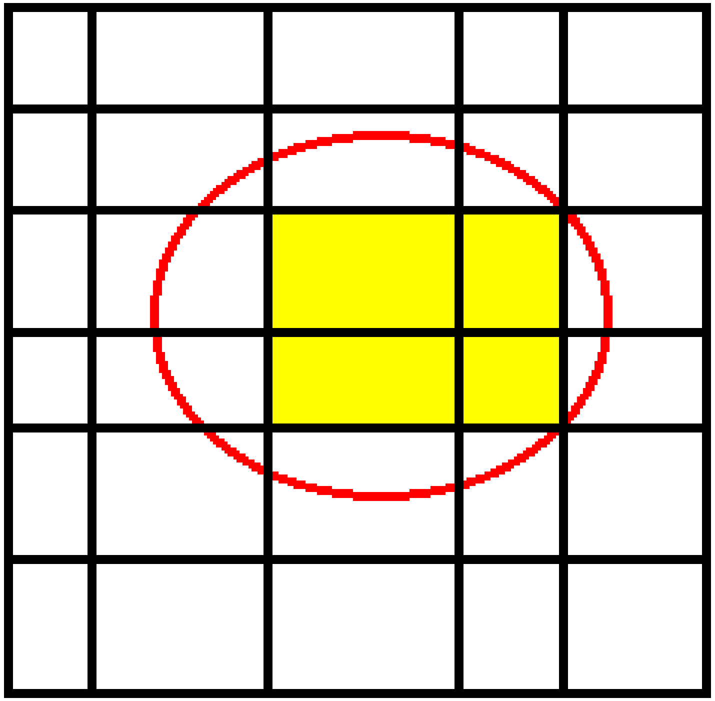
Podsumowanie

Obszar negatywny



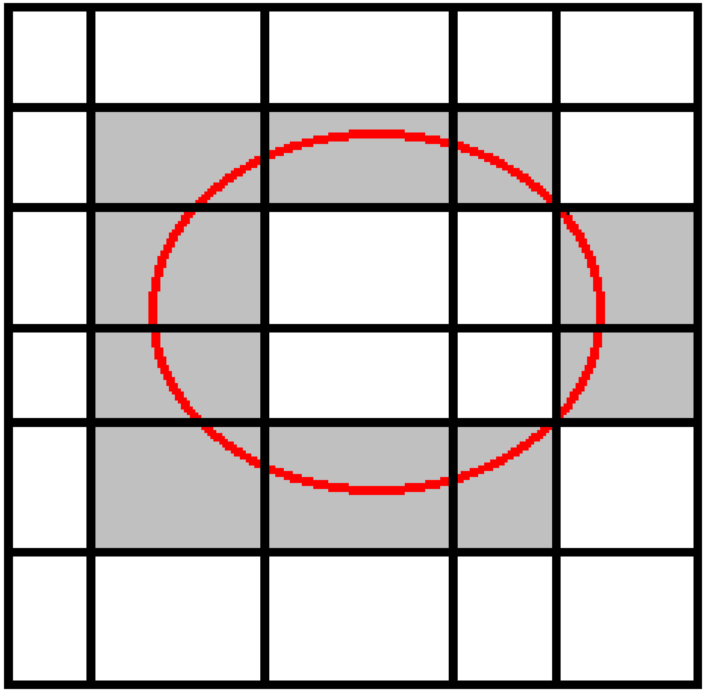
Podsumowanie

Obszar pozytywny - dolne przybliżenie



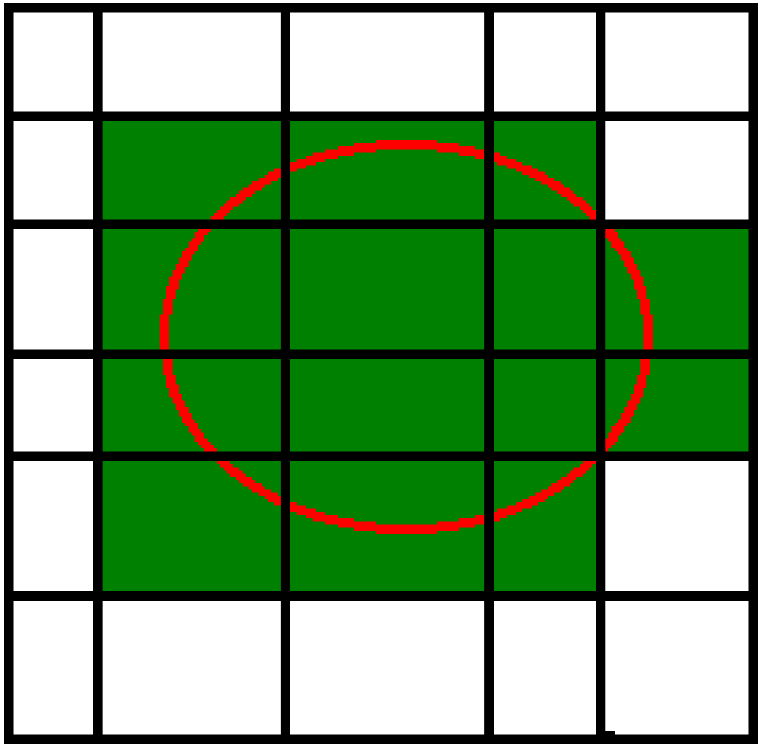
Podsumowanie

Obszar graniczny



Podsumowanie

Górne przybliżenie



Bibliografia

- **Zbiory przybliżone – wnioskowanie przybliżone** Autor: Nowotarski, Chodara ,Leończyk.
- **Metody wnioskowań aproksymacyjnych dla syntezy algorytmów decyzyjnych** Autor: Bazan J.
- **Rough Sets. International Journal of Komputer and Information Sciences** Autor:Pawlak Z.
- **Fuzzy sets, Information and Control**

  Autor: Zadeh, L., A.