Systemy uczące się – wykład 4

dr Przemysław Juszczuk

Katedra Inżynierii Wiedzy, Uniwersytet Ekonomiczny

XII 2018

Przykład

Miażdżyca powoduje często zwężenie tętnic wieńcowych. Prowadzi to zazwyczaj do zmniejszenia przepływu krwi w tych naczyniach, co może wywołać niedotlenienie mięśnia sercowego, zwłaszcza przy wysiłku fizycznym.

- Które fragmenty wskazują na niepewność wnioskowania?
- Jak przekształcić powyższy tekst w taki sposób, aby można było do niego zastosować jedną z poznanych dotychczas reprezentacji wiedzy.
- Czy do tak przekształconego tekstu można zadać pytania:
  - jaki ma wpływ wysiłek fizyczny na niedotlenienie mięśnia sercowego u ludzi z jednakowo posuniętą miażdżycą, wykonujących wysiłek fizyczny o różnym natężeniu?
  - w jakim stopniu człowiek, u którego nie występuje niedotlenienie z powodu wysiłku, narażony jest na zwężenie tętnic z powodu miażdżycy?

Definicje

Podejście probabilistyczne:
Mając dany zbiór hipotez:

$$H = \{h_1, ..., h_n\}$$

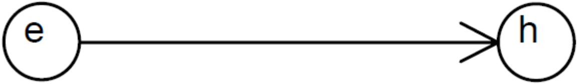
dla których:

$$P(h_i) > 0 \text{ dla każdego } i$$

Mając zbiór pewnych obserwacji:

$$E = \{e_1, ..., e_m\}$$

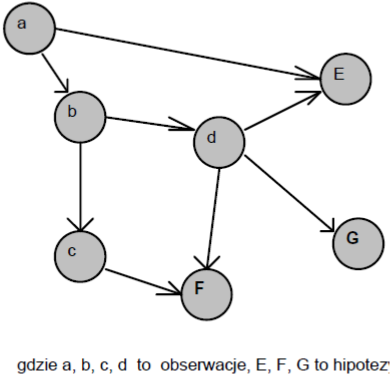
każdy fragment obserwacji e_j jest niezależny warunkowo względem każdej hipotezy.



Rysunek: Przesłanka a hipoteza

Obserwacja e oraz hipoteza h są reprezentowane przez wierzchołki grafu, natomiast natomiast wnioskowanie przez krawędź. Rozpatrywana reguła może być rozpatrywana w modelu Bayesa następująco:

$$P(h|e) = \frac{P(e|h) \cdot P(h)}{P(e)}$$



Rysunek: Sieć wnioskowania

Sieć Bayesa

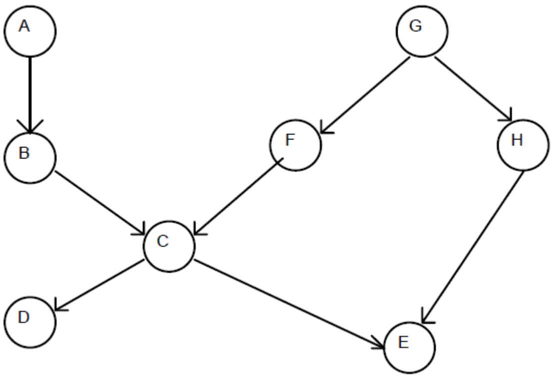
$$B = \{N, E, CP\}$$

gdzie dwójka {N, E} jest skierowanym grafem acyklicznym zbudowanym na podstawie zadanych prawdopodobieństw warunkowych zawartych w zbiorze CP. Przykład:
Dany jest zbiór pewnych zmiennych identyfikujących obserwacje i hipotezy. P Niech zbiór tych zmiennych ma następującą postać:

$$Z = A, B, C, D, E, F, G, H$$

Dane są również informacje opisujące związki przyczynowo-skutkowe pomiędzy tymi zmiennymi w postaci zbiory prawdopodobieństw warunkowych CP:

$$CP = \{P(A), P(B|A), P(C|B), P(C|F), P(D|C), P(E|CH), P(F|G), P(G), P(H|G)\}$$



Rysunek: Wynikowa sieć Bayesa

Współczynniki pewności

CF

Współczynniki pewności CF:

Jeżeli $e(\&e2\&...\&e_n)$ To h ze stopniem pewności CF

gdzie:

$e(, e2, ..., e_n)$ to przesłanki reguły a h to konkluzja, & to operator logiczny And.

Współczynnik CF

- CF nie jest interpretowany jako klasyczne prawdopodobieństwo
- Współczynnik pewności CF jest połączeniem stopnia wiedzy, oraz niewiedzy.
- Stopień wiedzy - inaczej miara wiarygodności - MB.
- Stopień niewiedzy - miara niewiarygodności - MD.

Załóżmy istnienie prostej reguły:

Jeżeli e to h

Współczynniki CF

Współczynniki dla powyższej reguły określone są następująco:

- MB(h,e)
- MD(h,e)
- CF(h,e)

Sam współczynnik CF definiowany jest jako:

$$CF(h, e) = MB(h, e) - MD(h, e)$$

Miary CF

Interpretacja powyższych miar może być następująca:

- jeżeli $P(h|e) = 1$ to h jest prawdziwe na pewno, wtedy $MB(h, e) = 1, MD(h, e) = 0$, oraz $CF(h, e) = 1$,
- jeżeli $P(\ h|e) = 1$ to h jest fałszywe na pewno, wtedy $MB(h, e) = 0, MD(h, e) = 1$, oraz $CF(h, e) = -1$,
- jeżeli $P(h|e) = P(h)$ to h co znaczy, że h i e są niezależne, wtedy $MB(h, e) = 0$, oraz $MD(h, e) = 0, CF(h, e) = 0$.

$$CF(h,e) = \begin{cases} 1, & P(h) = 1, \\ MB(h,e) & P(h|e) > P(h) \\ 0, & P(h|e) = P(h) \\ -MD(h,e), & P(h|e) < P(h) \\ -1, & P(h) = 0 \end{cases}$$

Rysunek: Wartości CF

Propagacja współczynników niepewności

Mając daną regułę R:

Jeżeli e to h ze stopniem pewności CF

- przesłanka reguły e ma pewien współczynnik CF
- konkluzja reguły h również ma współczynnik CF

Końcowy współczynnik pewności wyznaczany jest w następujący sposób:

$$CF(h, e) = CF(e) \cdot CF(h)$$

Współczynniki pewności

- W przypadku gdy przesłanka reguły zawiera wyrażenie zawierające operator AND (&) :

  Jeżeli $e1\&e2$ to h ze stopniem pewności CF

  to współczynnik pewności konkluzji h wyznaczany jest w następujący sposób:

  $$CF(h, e1\&e2) = \text{Minimum } \{CF(e1), CF(e2)\} \cdot CF(h)$$

- W przypadku gdy przesłanka reguły zawiera wyrażenie zawierajace funktor OR (|) :

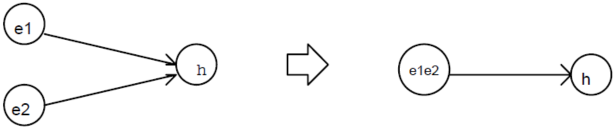
  Jeżeli $e1|e2$ to h ze stopniem pewności CF

  to współczynnik pewności konkluzji h wyznaczany jest w następujący sposób:

  $$CF(h, e1|e2) = \text{Maksimum } \{CF(e1), CF(e2)\} \cdot CF(h)$$

W przypadku, gdy jedna hipoteza h jest konkluzją więcej niż jednej reguły:

- Jeżeli e1 to h
- Jeżeli e2 to h



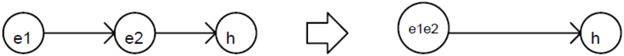
Rysunek: Obliczanie CF

$$CF(h,e_1,e_2)=\begin{cases} CF(h,e_1)+CF(h,e_2)-CF(h,e_1)CF(h,e_2), & CF(h,e_1)\ i\ CF(h,e_2)\geq 0 \\ \dfrac{CF(h,e_1)+CF(h,e_2)}{1-\min(|CF(h,e_1)|,|CF(h,e_2)|)}, & CF(h,e_1)*CF(h,e_2)<0, \\ CF(h,e_1)+CF(h,e_2)+CF(h,e_1)CF(h,e_2), & CF(h,e_1)\ i\ CF(h,e_2)<0 \end{cases}$$

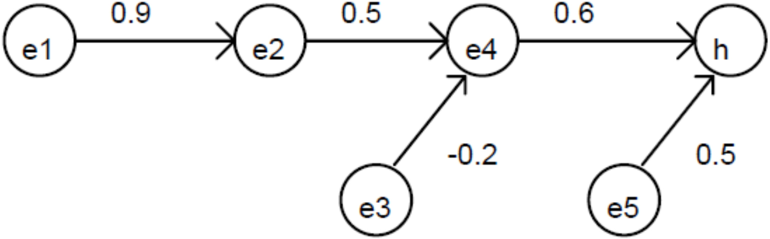
Rysunek: Obliczanie CF

Połączenie szeregowe reguł:

- Jeżeli e1 to e2
- Jeżeli e2 to h

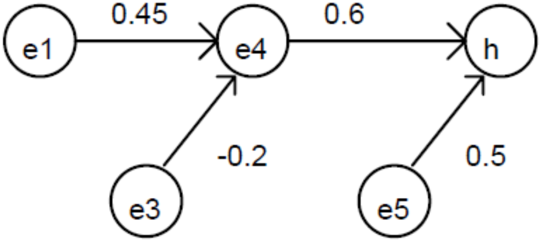


Rysunek: Szeregowe połączenie reguł $CF(h, e1) = CF(e2, e1) \cdot CF(h, e2)$



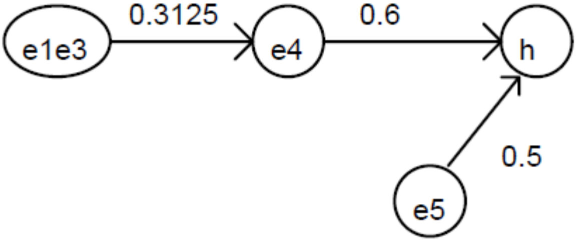
Rysunek: Propagacja CF

$$CF(e4, e1) = CF(e2, e1) \cdot CF(e4, e2)$$



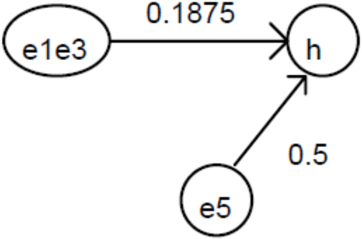
Rysunek: Propagacja CF

$$CF(e4, e1, e3) = \frac{CF(e4,e1)+CF(e4,e3)}{1-min(|CF(e4,e1)|,|CF(e4,e3)|)|}$$
$$CF(e4, e1, e3) = \frac{0.25}{1-0.2} = \frac{0.25}{0.8} = 0.3125$$



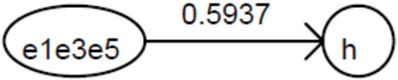
Rysunek: Propagacja CF

$CF(h, e4) = CF(e4, e1e3) \cdot CF(h, e1e3)$



Rysunek: Propagacja CF

$$CF(h, e1e3, e5) = CF(h, e1e3) + CF(h, e5) - CF(h, e1e3) \cdot CF(h, e5)$$



Rysunek: Propagacja CF

$CF(h, e1e3e5) = 0.5937$

Definicja

Sieć składająca się z następujących elementów:

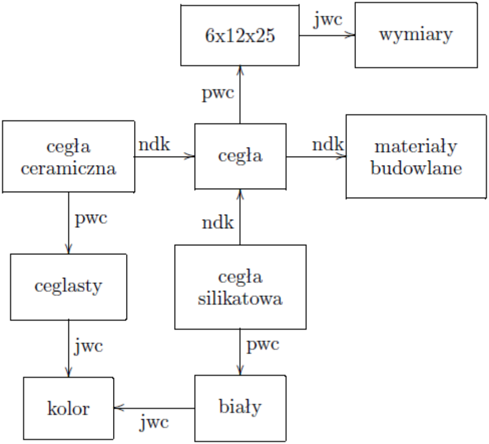
- zbiór obiektów $\{o_j\} = O$
- zbiór cech $\{c_j\} = C$
- zbiór wartości $\{v_j\} = V$

Elementami zbioru obiektów moga byc symbole oznaczajace konkrety lub abstrakcje, np:samochód - to symbol abstrakcji, zas Fiat 126 p KAE 0321 - to symbol konkretu.

Relacje

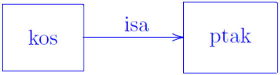
- *OxO* relacja miedzy obiektami, relacja okreslona na zbiorze obiektów.
  - ISA - relacja typu „cześć- całość” (nadrzędność)
  - ISPART - relacja podrzędności, czyli „(coś) jest cześcią (czegoś)”

  Relacje te są przechodnie.
- *OxC* - relacja przysługiwania obiektom pewnych cech - „posiada cechę”.
- *VxC* - relacja postaci „jest wartością cechy”.
- *VxV* - relacja uporządkowania elementów zbioru wartości cech, np.: relacja typu: „(coś) poprzedza (coś)”, lub „(coś) następuje po (czymś)”.
- *OxV* - relacja typu „posiada wartość cechy”, czyli przypisania obiektom wartości cechy. Czasem relacja ta jest tworzona przez złączenie relacji *OxC* oraz *VxC*.



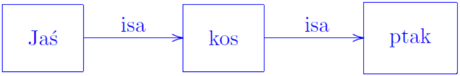
Rysunek: Sieć semantyczna - przykład

- Wszystkie kosy są ptakami:



Rysunek: Budowanie sieci semantycznej - przykład

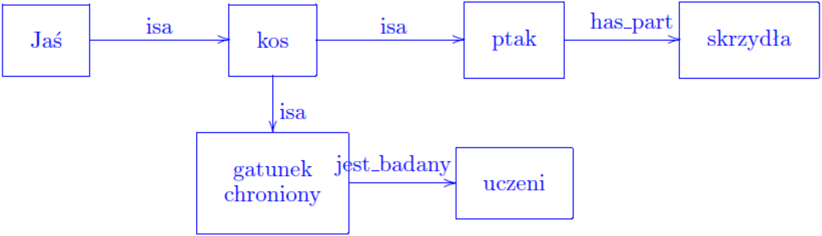
- Jaś jest kosem



Rysunek: Budowanie sieci semantycznej - przykład

Możliwe jest następujące wnioskowanie:

- Jaś jest kosem,
- kos jest ptakiem,
- Jaś jest ptakiem.

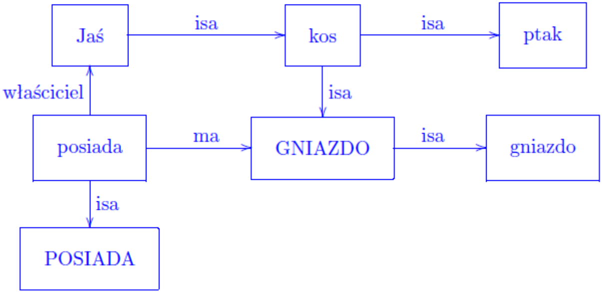


Rysunek: Sieć semantyczna - błąd wnioskowania

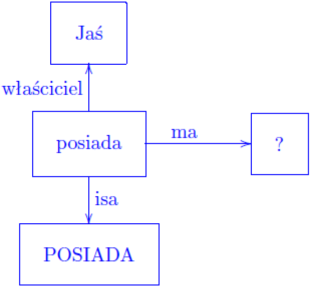
Ale: „uczeni badaja Jasia”, co może ale nie musi być prawdą.

Co posiada Jaś ?

System informacyjny



Rysunek: Sieć semantyczna - system informacyjny



Rysunek: Sieć semantyczna - zapytanie

Zastosowanie sieci semantycznych

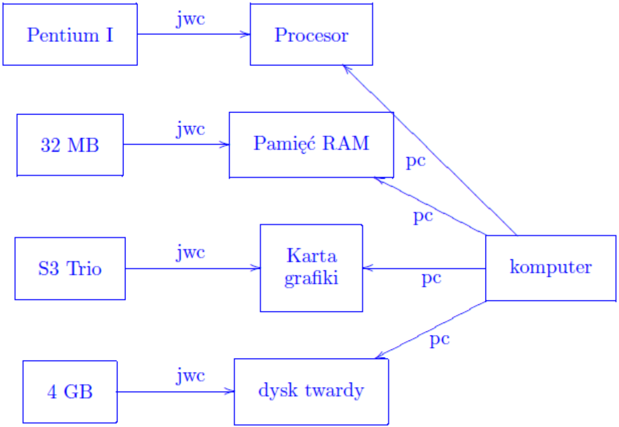
- projektowanie systemów informacyjnych (baz danych);
- rozumienie języka naturalnego;
- rozpoznawanie mowy;
- budowania systemów odpowiadajacych na pytania

Przykład 1

Komputer jest opisywany przez nastepujace parametry:

- procesor;
- pamięć RAM;
- karta grafiki;
- dysk twardy.

Przedstaw sieć semantyczną opisującą powyższe zależności. Jako przykład przyjmij komputer: procesor Pentium I, pamieć RAM 32 MB, karta grafiki S3 Trio, dysk HDD 4GB.



Rysunek: Rozwiązanie

Dziękuję za uwagę