

Modelowanie niepewności II

Sieci bayesowskie

Dr hab. inż. Joanna Józefowska, prof. PP

1/1

Plan wykładu

- Twierdzenie Bayesa
- Definicja sieci bayesowskiej
- Lokalna semantyka sieci bayesowskiej
- Globalna semantyka sieci bayesowskiej
- Wnioskowanie w sieciach bayesowskich
- Budowa sieci bayesowskich
- Podsumowanie

Dr hab. inż. Joanna Józefowska, prof. PP

1/2

Wnioskowanie probabilistyczne

B – włamanie

E – trzęsienie ziemi

A – alarm

J – sąsiad dzwoni

M – sąsiadka dzwoni

Łączny rozkład prawdopodobieństwa – $P(B, E, A, J, M)$

Dr hab. inż. Joanna Józefowska, prof. PP

1/3

B	E	A	J	M	P(B,E,A,J,M)
1	1	1	1	1	0.0001197000
1	1	1	1	0	0.0000133000
1	1	1	0	1	0.0000513000
1	1	1	0	0	0.0000057000
1	1	0	1	1	0.0000000050
1	1	0	1	0	0.0000000050
1	1	0	0	1	0.0000004950
1	1	0	0	0	0.0000009405
1	0	1	1	1	0.0058035600
1	0	1	1	0	0.0006448400
1	0	1	0	1	0.0024872400
1	0	1	0	0	0.0002763600
1	0	0	1	1	0.0000002940
1	0	0	1	0	0.0000005860
1	0	0	0	1	0.0000291060
1	0	0	0	0	0.0005530140
0	1	1	1	1	0.0036174600
0	1	1	1	0	0.0004019400
0	1	1	0	1	0.0015503400
0	1	1	0	0	0.0001722400
0	1	0	1	1	0.0000070290
0	1	0	1	0	0.0001335510
0	1	0	0	1	0.0006958710
0	1	0	0	0	0.0132215490
0	0	1	1	1	0.0006112260
0	0	1	1	0	0.0000679140
0	0	1	0	1	0.0002619540
0	0	1	0	0	0.0000291060
0	0	0	1	1	0.0004846149
0	0	0	1	0	0.0092076831
0	0	0	0	1	0.0479768751
0	0	0	0	0	0.9115606269

Łączny rozkład prawdopodobieństwa

Dr hab. inż. Joanna Józefowska, prof. PP

1/4

Wnioskowanie probabilistyczne

Jakie jest prawdopodobieństwo włamania, jeżeli zadzwoniła sąsiadka: $P(B=y|M=y)$?

Obliczamy prawdopodobieństwo brzegowe:

$$P(B, M) = \sum_{E, A, J} P(B, E, A, J, M)$$

B	M	P(B,E,A,J,M)
1	1	0.0084917
1	0	0.0015083
0	1	0.05520537
0	0	0.93479463

Obliczamy prawdopodobieństwo warunkowe:

$$P(B|M) = \frac{P(B, M)}{P(M)} = \frac{0.0084917}{0.0084917 + 0.05520537} = 0.133313$$

Dr hab. inż. Joanna Józefowska, prof. PP

1/5

Zalety wnioskowania probabilistycznego

- Dobrze zdefiniowana teoria matematyczna
- Na podstawie łącznego rozkładu prawdopodobieństwa można prowadzić dowolne wnioskowanie:
 - o przyczynach na podstawie obserwowanych skutków,
 - o skutkach na podstawie przesłanek
 - różne kombinacje powyższych
- Przejrzysta semantyka wynikająca z rachunku prawdopodobieństwa
- Model można uczyć z danych metodami statystycznymi

Dr hab. inż. Joanna Józefowska, prof. PP

1/6

Złożoność wnioskowania probabilistycznego

- w przykładzie z alarmem
 - trzeba znać $(2^5 - 1) = 31$ wartości,
 - mamy bezpośredni dostęp do mało przydatnych informacji np.

$$P(B=1, E=1, A=1, J=1, M=1)$$

- obliczenie praktycznej wartości $P(B=1|M=1)$ wymaga 29 dodawań

- w ogólności
 - $P(X_1, \dots, X_n)$ wymaga zapamiętania $2^n - 1$ wartości prawdopodobieństwa
 - trudne pozyskiwanie wiedzy (złożone, nienaturalne)
 - Wykładnicza złożoność pamięciowa i obliczeniowa

Dr hab. inż. Joanna Jozefowska, prof. P.P.

1/7

Twierdzenie Bayesa

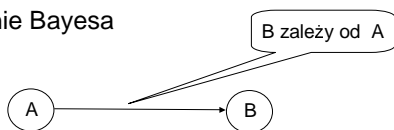
Jeśli uwzględnimy twierdzenie o prawdopodobieństwie całkowitym oraz użyjemy symboli H i E dla zdarzeń oznaczających odpowiednio hipotezę oraz obserwację (przesłankę), to reguła Bayesa przyjmuje postać:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

Dr hab. inż. Joanna Jozefowska, prof. P.P.

1/8

Twierdzenie Bayesa



$$P(B|A)$$

$$P(B) = \sum_A P(B|A)P(A)$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Dr hab. inż. Joanna Jozefowska, prof. P.P.

1/9

Reguła łańcucha

$$P(X_1, X_2) = P(X_1)P(X_2|X_1)$$

$$P(X_1, X_2, X_3) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)$$

$$\dots\dots\dots$$

$$P(X_1, X_2, \dots, X_n) = P(X_1)P(X_2|X_1) \dots P(X_n|X_1, \dots, X_{n-1})$$

Dr hab. inż. Joanna Jozefowska, prof. P.P.

1/10

Warunkowa niezależność zmiennych w dziedzinie

W danej dziedzinie zwykle można zdefiniować zbiór $pa(X_i) \subseteq \{X_1, \dots, X_{i-1}\}$ taki, że X_i jest niezależne od zmiennych ze zbioru $\{X_1, \dots, X_{i-1}\} \setminus pa(X_i)$.

Zatem

$$P(X_i|X_1, \dots, X_{i-1}) = P(X_i|pa(X_i))$$

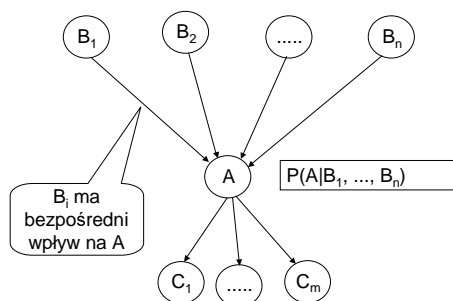
oraz

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i|pa(X_i))$$

Dr hab. inż. Joanna Jozefowska, prof. P.P.

1/11

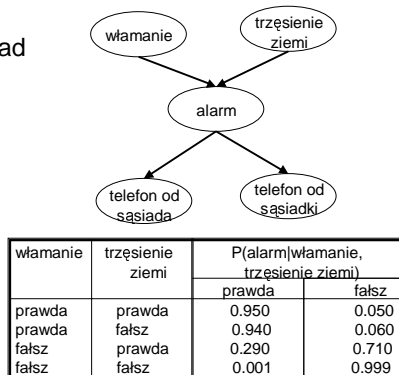
Sieć bayesowska



Dr hab. inż. Joanna Jozefowska, prof. P.P.

1/12

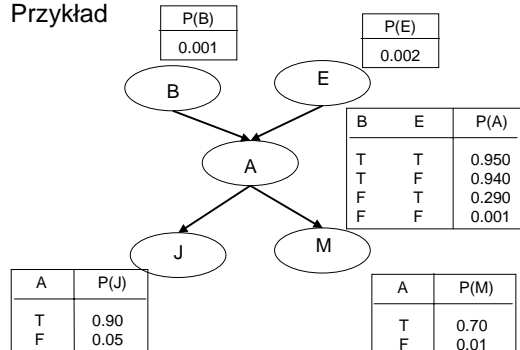
Przykład



Dr hab. inż. Joanna Jozefowska, prof. PP

1/13

Przykład



Dr hab. inż. Joanna Jozefowska, prof. PP

1/14

Złożoność reprezentacji

- Zamiast 31 wartości wystarczy znać 10.
- Model łatwiej zbudować
 - wystarczy mniejsza liczba parametrów.
 - parametry bardziej naturalnie wynikają z problemu
- Łatwiejsze wnioskowanie.

Dr hab. inż. Joanna Jozefowska, prof. PP

1/15

Sieci bayesowskie

Sieć bayesowska jest to acykliczny graf skierowany (DAG), którego

- wierzchołki reprezentują wyrażenia rachunku zdań lub zmienne losowe typowe dla danej dziedziny.
- łuki reprezentują relację bezpośredniej zależności zmiennych, które są kwantyfikowane w postaci odpowiednich rozkładów prawdopodobieństw warunkowych, wyrażających przekonanie o sile tych zależności.

Dr hab. inż. Joanna Jozefowska, prof. PP

1/16

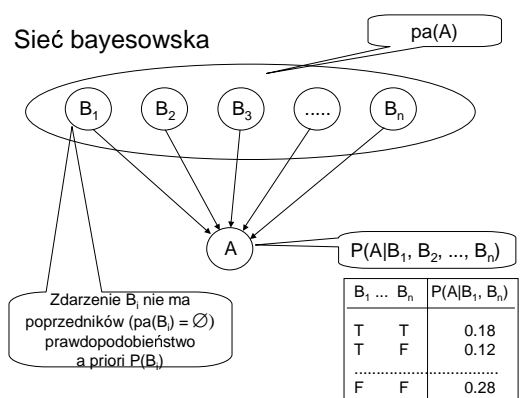
Sieci bayesowskie

Każdej zmiennej A posiadającej rodziców B_1, \dots, B_n przypisana jest tablica zawierająca rozkład prawdopodobieństwa warunkowego $P(A|B_1, \dots, B_n)$, lub w skrócie $P(A|pa(A))$ (gdzie $pa(A)$ oznacza zbiór wszystkich rodziców zmiennej A). W szczególnym przypadku, gdy $pa(A) = \emptyset$ mówimy o prawdopodobieństwie *a priori* równym $P(A)$. Struktura grafu odzwierciedla założenia o niezależności zmiennych występujących w problemie.

Dr hab. inż. Joanna Jozefowska, prof. PP

1/17

Sieć bayesowska



Dr hab. inż. Joanna Jozefowska, prof. PP

1/18

Lokalna semantyka sieci bayesowskiej

Wiedza o problemie jest przedstawiona w postaci zbioru bezpośrednich zależności między zmiennymi.

Zatem wystarczy podać prawdopodobieństwo warunkowe zmiennej względem jej bezpośrednich poprzedników (rodziców).

Zakłada się równocześnie, że jeżeli znamy wartości rodziców danej zmiennej, to jej wartość nie zależy od wartości pozostałych zmiennych w sieci (warunkowa niezależność zmiennych).

Dr hab. inż. Joanna Jozefowska, prof. PP

1/19

Globalna semantyka sieci bayesowskiej

Globalnie sieć bayesowska reprezentuje łączny rozkład prawdopodobieństwa występujących w niej zmiennych. Łączny rozkład prawdopodobieństwa dany jest w sposób niejawni. Rozkład ten wyznacza się zgodnie z regułą łańcucha w sposób następujący:

$$P(A_1, \dots, A_n) = \prod_i P(A_i | A_{i+1}, \dots, A_n) \\ = P(A_1 | A_2, A_3, \dots, A_n) \dots P(A_{n-1} | A_n) P(A_n)$$

Dr hab. inż. Joanna Jozefowska, prof. PP

1/20

Globalna semantyka sieci bayesowskiej

Numerując wierzchołki grafu tak by indeks przypisany zmiennej był mniejszy niż indeksy przypisane jego przodkom oraz na mocy warunkowej niezależności otrzymujemy:

$$P(A_i | A_{i+1}, \dots, A_n) = P(A_i | pa(A_i))$$

Ostatecznie mamy:

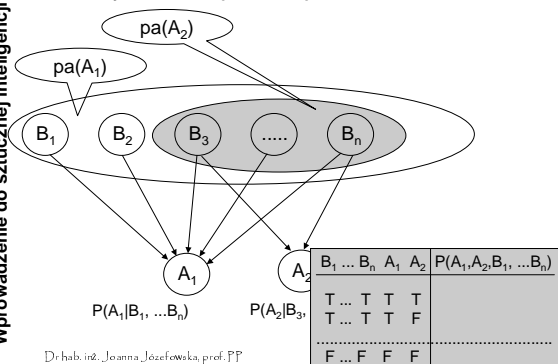
$$P(A_1, \dots, A_n) = \prod_i P(A_i | pa(A_i))$$

Sieć bayesowska jest więc pełnym modelem probabilistycznym.

Dr hab. inż. Joanna Jozefowska, prof. PP

1/21

Globalny rozkład prawdopodobieństwa



Dr hab. inż. Joanna Jozefowska, prof. PP

Wnioskowanie w sieciach bayesowskich

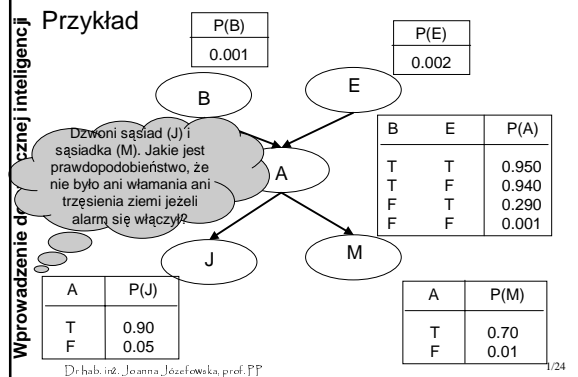
Wnioskowanie probabilistyczne (także w sieciach bayesowskich) polega na aktualizacji przekonania o prawdziwości hipotez w przypadku wystąpienia pewnych obserwacji (przesłanek), czyli na wyznaczeniu rozkładu prawdopodobieństwa warunkowego $P(H|E)$ dla hipotezy H , gdzie E oznacza zbiór obserwacji.

Niezależnie od struktury sieci bayesowskiej rozróżnia się przypadek szczególny wnioskowania (wyznaczenie rozkładu prawdopodobieństwa dla pojedynczej hipotezy) i ogólny (wyznaczenie rozkładu prawdopodobieństwa dla wszystkich hipotez).

Dr hab. inż. Joanna Jozefowska, prof. PP

1/23

Przykład



Dr hab. inż. Joanna Jozefowska, prof. PP

1/24

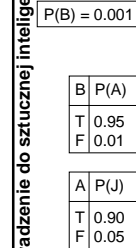
$$P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) = ?$$

$$\begin{aligned} P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) \\ = P(J|A)P(M|A)P(A|\neg B \wedge \neg E)P(\neg B)P(\neg E) \\ = 0.9 * 0.7 * 0.001 * 0.999 * 0.998 \\ = 0.00062 \end{aligned}$$

Dr hab. inż. Joanna Jozefowska, prof. PP

1/25

Rodzaje wnioskowania w sieciach bayesowskich



Wnioskowanie diagnostyczne

Pojawia się przesłanka J i chcemy zaktualizować przekonanie o prawdziwości hipotezy B.

Innymi słowy widzimy objawy i szukamy przyczyny (diagnozy).

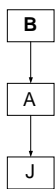
Sąsiad dzwoni, jakie jest prawdopodobieństwo, że włamano się do domu?

$$P(B|J) = 0.016$$

Dr hab. inż. Joanna Jozefowska, prof. PP

1/26

Rodzaje wnioskowania w sieciach bayesowskich



b)
przyczynowe

Pojawia się przesłanka B i chcemy zaktualizować przekonanie o prawdziwości hipotezy J.

Innymi słowy znamy przyczynę i szukamy możliwych skutków (interpretacja).

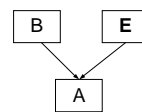
Włamano się do domu. Jakie jest prawdopodobieństwo, że sąsiad zadzwoni?

$$P(J|B) = P(J|A)P(A|B) = 0.9 * 0.95 = 0.86$$

Dr hab. inż. Joanna Jozefowska, prof. PP

1/27

Rodzaje wnioskowania w sieciach bayesowskich



c)
między-przyczynowe

Pojawia się przesłanka E i chcemy zaktualizować przekonanie o prawdziwości hipotezy B.

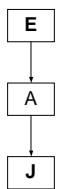
Innymi słowy wystąpienie innej przyczyny może wpłynąć na prawdopodobieństwo uznania drugiej przyczyny.

Włączył się alarm, więc $P(B|A) = 0.376$, ale jeżeli równocześnie wystąpi trzęsienie ziemi, to $P(B|A, E) = 0.03$

Dr hab. inż. Joanna Jozefowska, prof. PP

1/28

Rodzaje wnioskowania w sieciach bayesowskich



d)
mieszane

Pojawiają się przesłanki E i J i chcemy zaktualizować przekonanie o prawdziwości hipotezy A.

Innymi słowy znamy przyczynę i skutek, szukamy prawdopodobieństwa zdarzenia pośredniego.

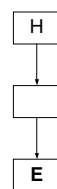
Sąsiad dzwoni i wiemy, że było trzęsienie ziemi. Jakie jest prawdopodobieństwo, że alarm się włączył?

$$P(A|J, E) = 0.03$$

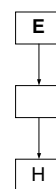
Dr hab. inż. Joanna Jozefowska, prof. PP

1/29

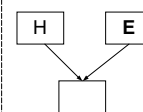
Rodzaje wnioskowania w sieciach bayesowskich



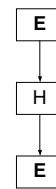
a)
diagnostyczne



b)
przyczynowe



c)
między-przyczynowe

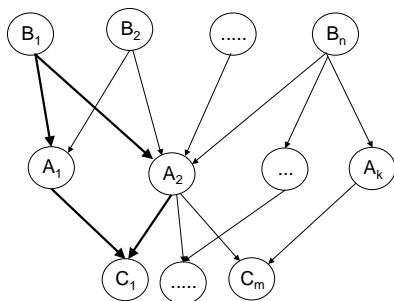


d)
mieszane

Dr hab. inż. Joanna Jozefowska, prof. PP

1/30

Sieć bayesowska wielokrotnie połączona



Dr hab. inż. Joanna Jozefowska, prof. PŁ

1/31

Podsumowanie

- Sieci bayesowskie stanowią metodę wnioskowania z zastosowaniem twierdzenia Bayesa redukującą złożoność obliczeń
- Wnioskowanie w sieciach bayesowskich jest w ogólności NP-trudne. W sieciach typu polidrzewa (single connected networks) wnioskowanie ma złożoność liniową.
- Powyższy problem rozwiązuje się na kilka sposobów:
 - Buduje się sieci od razu w postaci polidrzewa
 - Przekształca się sieć do postaci polidrzewa (drzewa złączy)
 - Stosuje się wnioskowanie przybliżone

Dr hab. inż. Joanna Jozefowska, prof. PŁ

1/32