

Modelowanie niepewności II

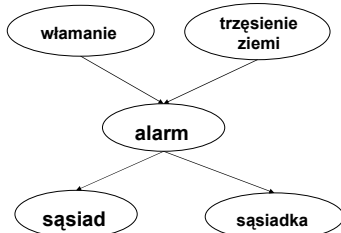


Sieci bayesowskie

Plan wykładu

- ▶ Twierdzenie Bayesa
- ▶ Definicja sieci bayesowskiej
- ▶ Lokalna semantyka sieci bayesowskiej
- ▶ Globalna semantyka sieci bayesowskiej
- ▶ Wnioskowanie w sieciach bayesowskich
- ▶ Narzędzia do projektowania SB
- ▶ Budowa sieci bayesowskich
- ▶ Podsumowanie

Wnioskowanie probabilistyczne



Wnioskowanie probabilistyczne

- B – włamanie
- E – trzęsienie ziemi
- A – alarm
- J – sąsiad dzwoni
- M – sąsiadka dzwoni



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

Wnioskowanie probabilistyczne

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

Obliczamy prawdopodobieństwo brzegowe:

B	M	P(B,M)
1	1	0,000659
1	0	0,000341
0	1	0,011078
0	0	0,987922

Obliczamy prawdopodobieństwo warunkowe:

$$P(B|M) = \frac{P(M,B)}{P(M)} = \frac{0,000659}{0,000659 + 0,011078} = 0,05611$$

prawdopodobieństwa

B	E	A	J	M	
1	1	1	1	1	0,0000011970
1	1	1	1	0	0,0000000130
1	1	1	0	1	0,0000000130
1	1	1	0	0	0,0000000070
1	1	0	1	1	0,0000000030
1	1	0	1	0	0,0000000020
1	1	0	0	1	0,0000000070
1	1	0	0	0	0,0000000030
1	0	1	1	1	0,0005973030
1	0	1	1	0	0,0002599870
1	0	1	0	1	0,0000660370
1	0	1	0	0	0,0000284430
1	0	0	1	1	0,0000314370
1	0	0	1	0	0,0000134730
1	0	0	0	1	0,0000034930
1	0	0	0	0	0,0000014970
0	1	1	1	1	0,0011958030
0	1	1	1	0	0,0005124870
0	1	1	0	1	0,0001328870
0	1	1	0	0	0,0000569930
0	1	0	1	1	0,0000639730
0	1	0	1	0	0,0000268730
0	1	0	0	1	0,0000069930
0	1	0	0	0	0,0000022970
0	0	1	1	1	0,9967066070
0	0	1	1	0	0,2557310130
0	0	1	0	1	0,0663006130
0	0	1	0	0	0,0284145570
0	0	0	1	1	0,0314056530
0	0	0	1	0	0,0134695270
0	0	0	0	1	0,0034895270

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

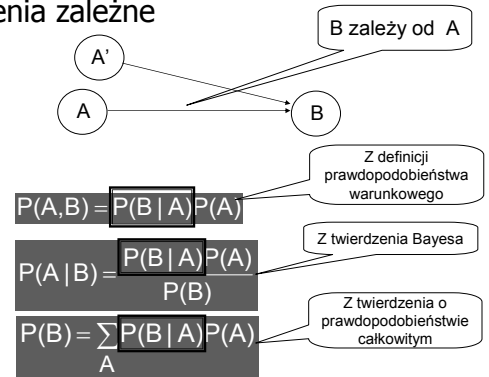
Złożoność wnioskowania probabilistycznego

- 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 prawdopodobieństw
 - trudne pozyskiwanie wiedzy (złożone, nienaturalne)
 - wykładnicza złożoność pamięciowa i obliczeniowa

Twierdzenie Bayesa

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

Zdarzenia zależne



Reguła łańcucha



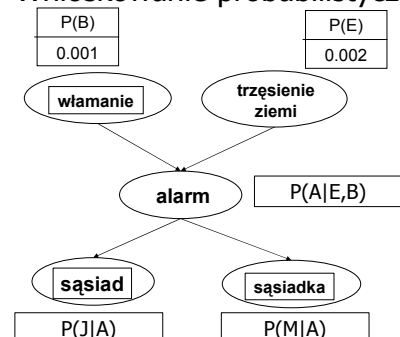
$$P(X_1, X_2) = P(X_1)P(X_2|X_1)$$

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

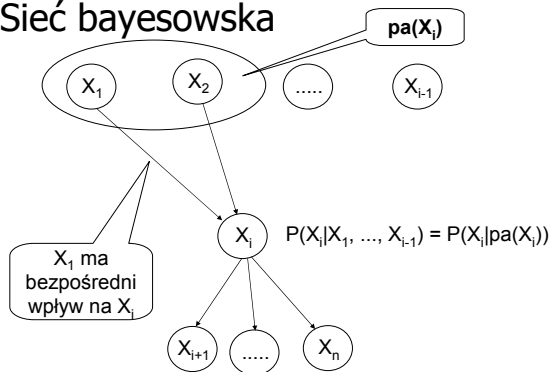
.....

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

Wnioskowanie probabilistyczne



Sieć bayesowska



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 $\{X_1, \dots, X_{i-1}\} \setminus pa(X_i)$.

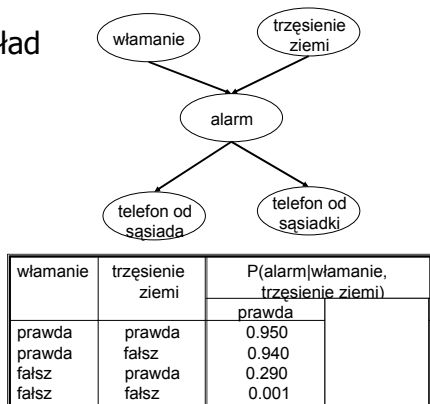
Zatem

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

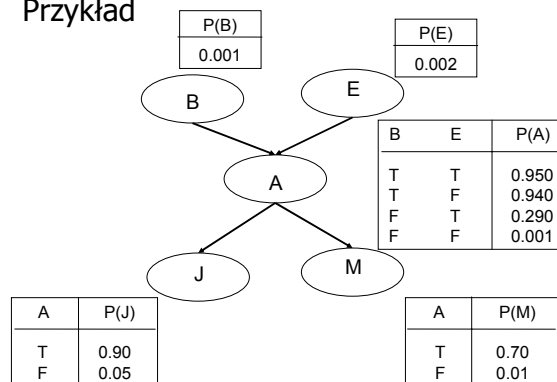
oraz

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

Przykład



Przykład



Złożoność reprezentacji

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

Sieci bayesowskie

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

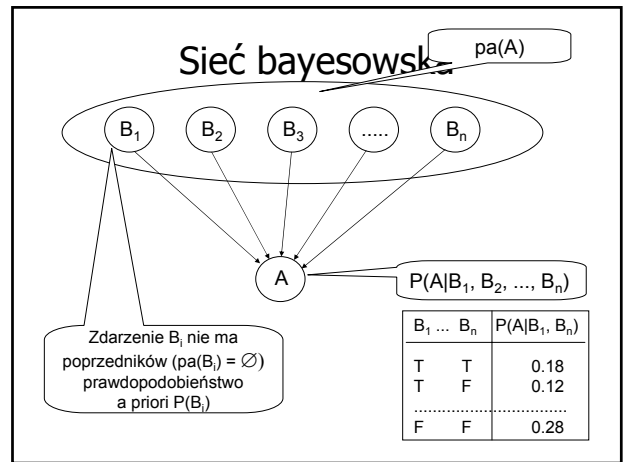
- wierzchołki reprezentują wyrażenia rachunku zdań lub zmienne losowe typowe dla danej dziedziny.
- łuki reprezentują relacje bezpośredniej zależności zmiennych, które są kwantyfikowane w postaci odpowiednich rozkładów prawdopodobieństw warunkowych, wyrażających przekonanie o sile tych zależności.

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.



Lokalna semantyka sieci bayesowskiej

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

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).

Globalna semantyka sieci bayesowskiej

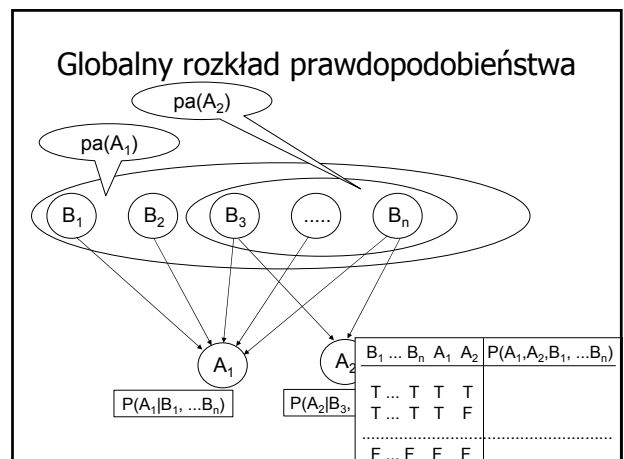
Globalnie sieć bayesowska reprezentuje łączny rozkład prawdopodobieństwa występujących w niej zmiennych. Łączny rozkład prawdopodobieństwa dany jest w sposób niejawnym. Rozkład ten wyznacza się zgodnie z regułą łańcucha w sposób następujący:

Globalna semantyka sieci bayesowskiej

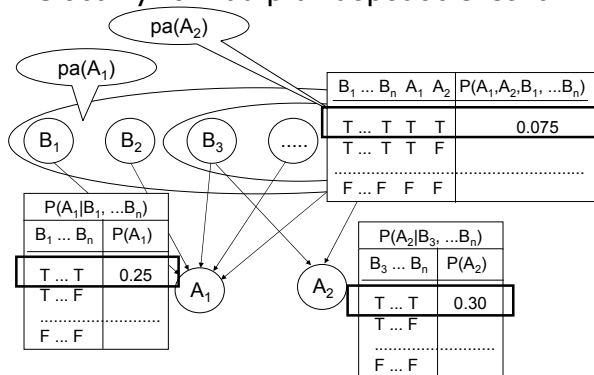
Numerując wierzchołki grafu tak by indeks przypisany zmiennej był mniejszy niż indeksy przypisane jego przodkom oraz na mocy warunkowej niezależności otrzymujemy:

Ostatecznie mamy:

Sieć bayesowska jest więc pełnym modelem probabilistycznym.



Globalny rozkład prawdopodobieństwa

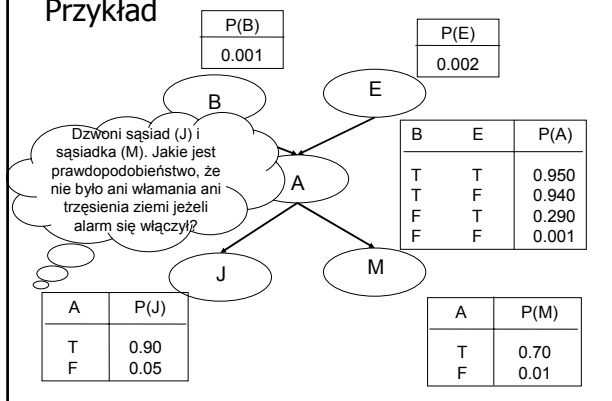


Wnioskowanie w sieciach bayesowskich

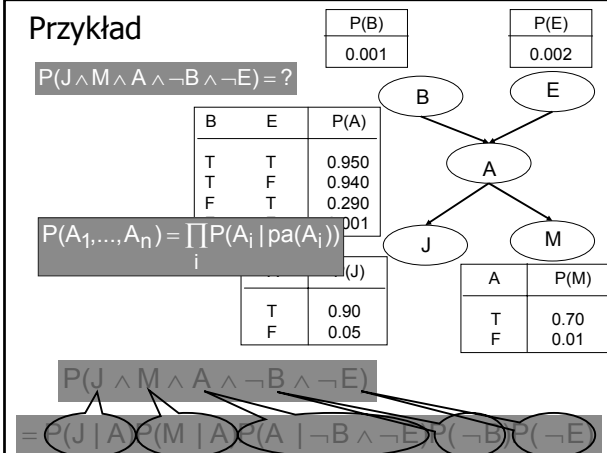
Wnioskowanie probabilistyczne (także w sieciach bayesowskich) polega na aktualizacji przekonania o prawdziwości hipotez w przypadku wystąpienia pewnych obserwacji (przesłanek), czyli na wyznaczeniu rozkładu prawdopodobieństwa warunkowego $P(H|E)$ dla hipotezy H , gdzie E oznacza zbiór obserwacji.

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

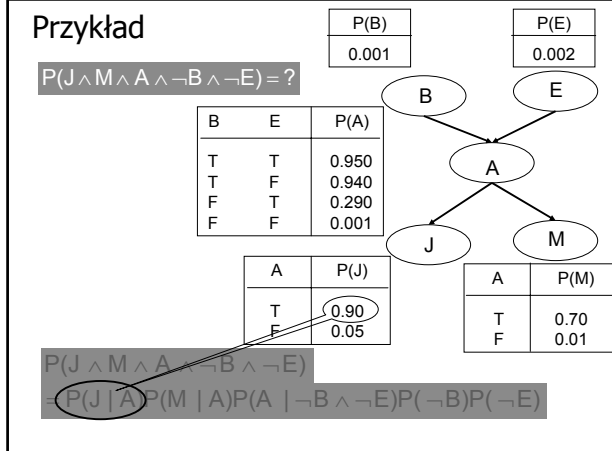
Przykład



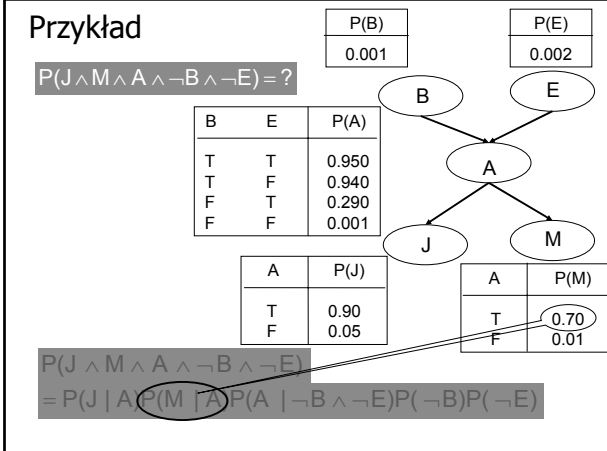
Przykład



Przykład

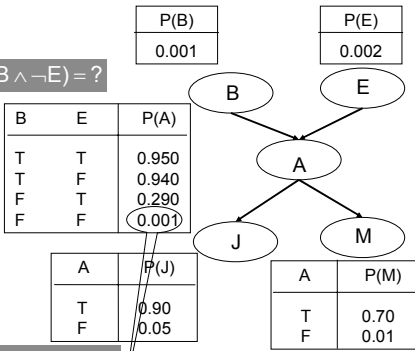


Przykład



Przykład

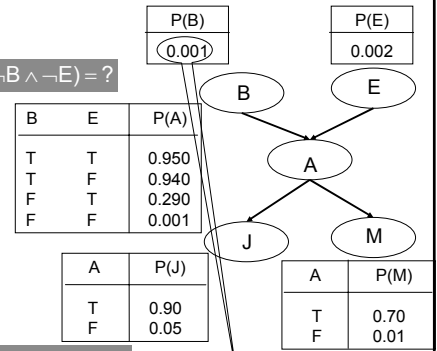
$$P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) = ?$$



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

Przykład

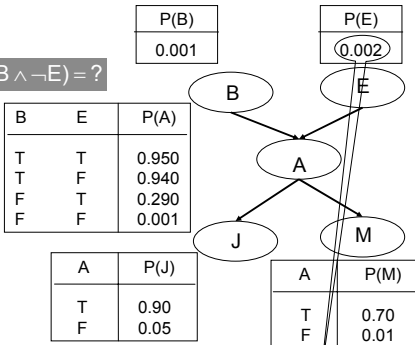
$$P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) = ?$$



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

Przykład

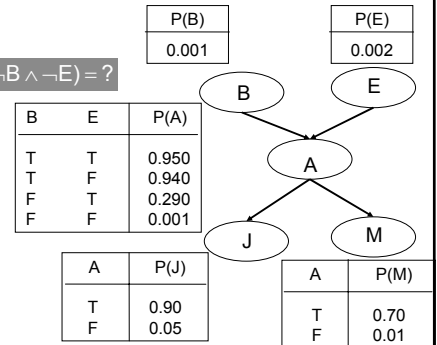
$$P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) = ?$$



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

Przykład

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



$$P(J \wedge M \wedge A \wedge \neg B \wedge \neg E) = P(J | A)P(M | A)P(A | \neg B \wedge \neg E)P(\neg B)P(\neg E) = 0.9 * 0.7 * 0.001 * 0.999 * 0.998 = 0.000628$$

Rodzaje wnioskowania w sieciach bayesowskich

Wnioskowanie przyczynowe

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

Innymi słowy znamy przyczynę i szukamy skutku.

$P(A|B) = P(A|B, E)P(E) + P(A|B, \neg E)P(\neg E) = 0.95 * 0.02 + 0.94 * 0.998 = 0.94002$

$P(J|B) = \frac{P(J|A)P(A|B)P(B) + P(J|\neg A)P(\neg A|B)P(B)}{P(B)} = \frac{0.9 * 0.94 + 0.05 * 0.06}{0.001} = 0.849$

przyczynowe

Rodzaje wnioskowania w sieciach bayesowskich

Wnioskowanie diagnostyczne

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

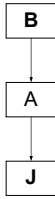
Innymi słowy widzimy objawy i szukamy przyczyny (diagnozy).

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

$P(B|J) = \frac{P(J|B)P(B)}{P(J)} = \frac{0.849 * 0.01}{0.05213} = 0.01628$

diagnostyczne

Rodzaje wnioskowania w sieciach bayesowskich



mieszane

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

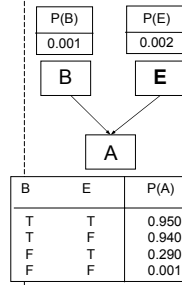
Innymi słowy znamy przyczynę i skutek, szukamy prawdopodobieństwa zdarzenia pośredniego.

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

$$P(A | J, B) = \frac{P(J | A)P(A | B)}{P(J | A)P(A | B) + P(J | \neg A)P(\neg A | B)}$$

$$= \frac{0,9 * 0,94}{0,9 * 0,94 + 0,05 * 0,06} = 0,9964$$

Rodzaje wnioskowania w sieciach bayesowskich



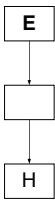
między-przyczynowe

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

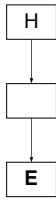
Innymi słowy wystąpienie innej przyczyny może wpłynąć na prawdopodobieństwo uznania drugiej przyczyny.

Włączył się alarm, więc $P(B|A) = 0.3736$, ale jeżeli równocześnie wystąpi trzęsienie ziemi, to $P(B|A, E) = 0.0032$

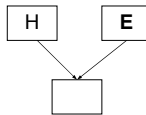
Rodzaje wnioskowania w sieciach bayesowskich



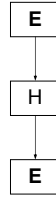
a)
przyczynowe



b)
diagnostyczne



c)
między-przyczynowe

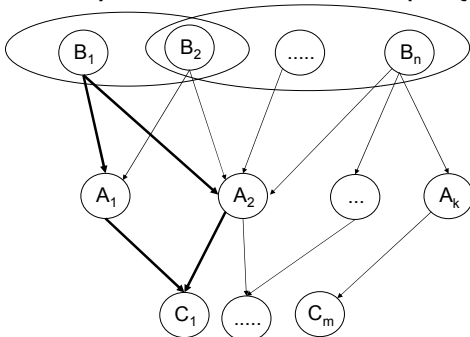


d)
mieszane

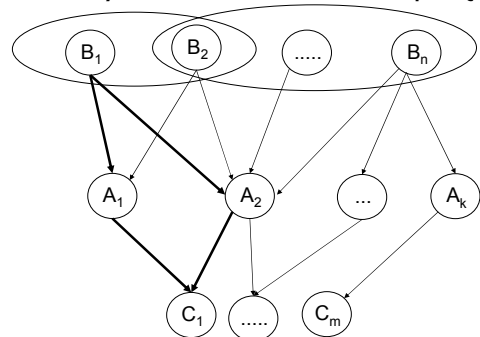
Narzędzia do projektowania sieci bayesowskich



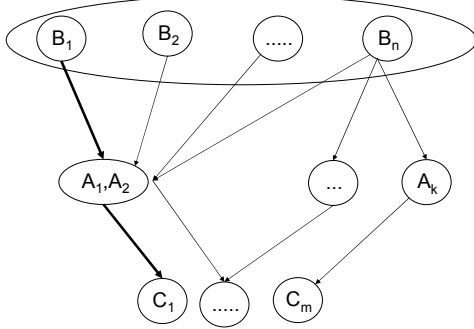
Sieć bayesowska wielokrotnie połączona



Sieć bayesowska wielokrotnie połączona



Sieć bayesowska wielokrotnie połączona



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