W sieci Bayesa

Łukasz Rajkowski

67. Szkoła Matematyki Poglądowej
Siedlce, 23 sierpnia 2024

Co to jest sieć Bayesa?

Co to jest sieć ~~Bayesa?~~ bayesowska?

Co to jest sieć ~~Bayesa?~~bayesowska?

60%

40%

Co to jest sieć ~~Bayesa?~~bayesowska?

Co to jest sieć ~~Bayesa?~~bayesowska?

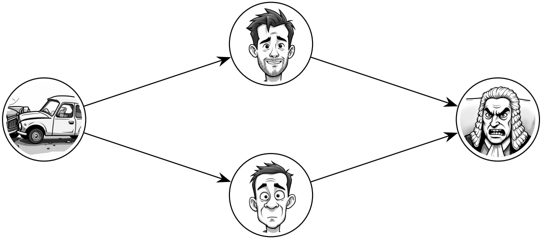
Co to jest sieć ~~Bayesa?~~bayesowska?

90%

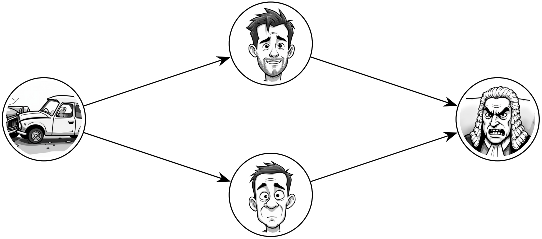
70%

Co to jest sieć ~~Bayesa?~~bayesowska?

Co to jest sieć ~~Bayesa?~~bayesowska?

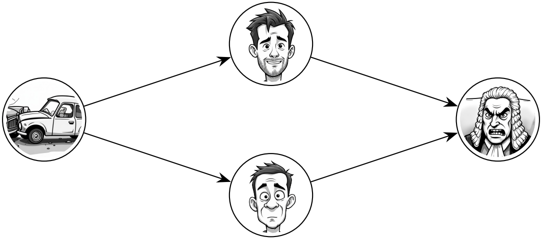


Co to jest sieć ~~Bayesa?~~bayesowska?



Sieć bayesowska to graf skierowany ilustrujący bezpośrednie zależności między zmiennymi.

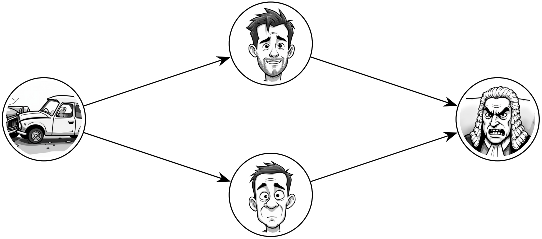
Co to jest sieć ~~Bayesa?~~bayesowska?



Sieć bayesowska to graf skierowany ilustrujący bezpośrednie zależności między zmiennymi. U nas:

- zeznania świadków mają bezpośredni wpływ na werdykt

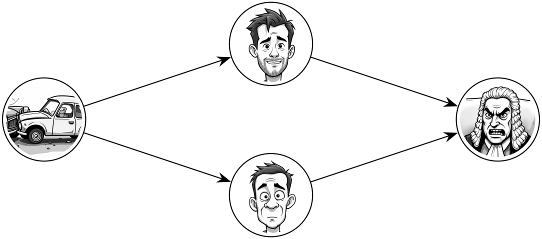
Co to jest sieć ~~Bayesa?~~bayesowska?



Sieć bayesowska to graf skierowany ilustrujący bezpośrednie zależności między zmiennymi. U nas:

- zeznania świadków mają bezpośredni wpływ na werdykt
- to, jaka taksówka spowodowała wypadek, ma oczywiście wpływ na werdykt, jednak nie bezpośredni

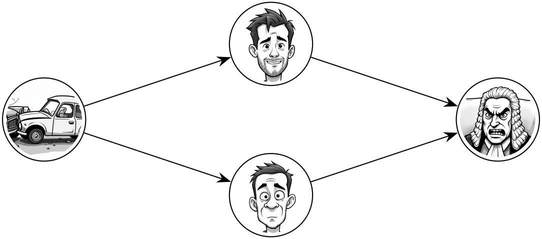
Co to jest sieć ~~Bayesa?~~bayesowska?



Sieć bayesowska to graf skierowany ilustrujący bezpośrednie zależności między zmiennymi. U nas:

- zeznania świadków mają bezpośredni wpływ na werdykt
- to, jaka taksówka spowodowała wypadek, ma oczywiście wpływ na werdykt, jednak nie bezpośredni
  jeśli znane są zeznania, prawda jest nieistotna dla werdyktu

Co to jest sieć ~~Bayesa?~~bayesowska?



Sieć bayesowska to graf skierowany ilustrujący bezpośrednie zależności między zmiennymi. U nas:

- zeznania świadków mają bezpośredni wpływ na werdykt
- to, jaka taksówka spowodowała wypadek, ma oczywiście wpływ na werdykt, jednak nie bezpośredni
  jeśli znane są zeznania, prawda jest nieistotna dla werdyktu

Formalna definicja już wkrótce!

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

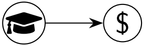
- **Dwie zmienne:**

(a) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

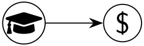
(a) 

(b) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

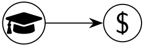
  (a) 

  (b) 

  Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne)

Przykłady małych sieci bayesowskich

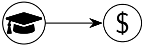
- **Dwie zmienne:**

  (a)  (b) 

  Piszemy $\scorpio \perp\!\!\!\perp \$$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp \$$ (zależne)

Przykłady małych sieci bayesowskich

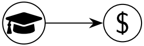
- **Dwie zmienne:**

  (a)  (b) 

  Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ \$ (zależne)
- **Trzy zmienne:**

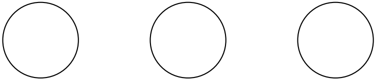
Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a)  (b) 

Piszemy ♏ $\perp\!\!\!\perp \$$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp \$$ (zależne)

- **Trzy zmienne:**

(a) 

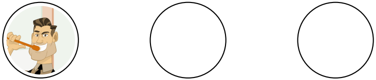
Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

  (a) ♏ $ (b) 🎓 → $

  Piszemy ♏ $\perp\!\!\!\perp$ $ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ $ (zależne)

- **Trzy zmienne:**

  (a) 

- **Dwie zmienne:**

  (a) ♏ $ (b) 🎓 → $

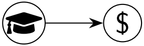
  Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\!\perp\!\!\!\perp$ \$ (zależne)

- **Trzy zmienne:**

  (a) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a)  (b) 

Piszemy ♏ $\perp\!\!\!\perp \$$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp \$$ (zależne)

- **Trzy zmienne:**

(a) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

  (a) ♏ $\qquad$ \$ $\qquad\qquad$ (b) 🎓 $\longrightarrow$ \$

  Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ \$ (zależne)

- **Trzy zmienne:**

  (a) 

  Mamy  $\not\perp\!\!\!\perp$ 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

  (a) ♏ $ (b) 🎓 → $

  Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ \$ (zależne)

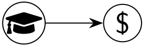
- **Trzy zmienne:**

  (a) 

  Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$ 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a)  (b) 

Piszemy ♏ $\perp\!\!\!\perp$ $\$$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ $\$$ (zależne)

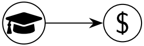
- **Trzy zmienne:**

(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$  (warunkowa niezależność)

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

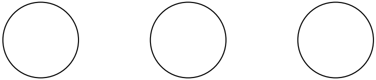
(a)  (b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz  $\not\perp\!\!\!\perp$ \$ (zależne)

- **Trzy zmienne:**

(a) 

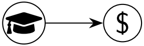
Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  |  (warunkowa niezależność)

(b) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a) 

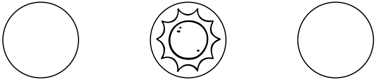
(b) 

Piszemy ♏ $\perp\!\!\!\perp$ $\$$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ $\$$ (zależne)

- **Trzy zmienne:**

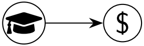
(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$  (warunkowa niezależność)

(b) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

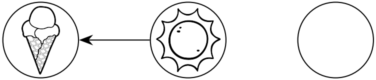
(a)  (b) 

Piszemy ♏ $\perp\!\!\!\perp$ $\$$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ $\$$ (zależne)

- **Trzy zmienne:**

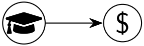
(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$  (warunkowa niezależność)

(b) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

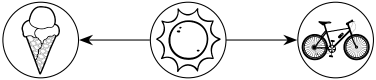
(a)  (b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ \$ (zależne)

- **Trzy zmienne:**

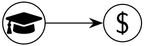
(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  |  (warunkowa niezależność)

(b) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

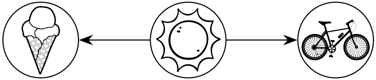
(a) 
(b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz  $\not\perp\!\!\!\perp$ \$ (zależne)

- **Trzy zmienne:**

(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$  (warunkowa niezależność)

(b) 

Tutaj  $\not\perp\!\!\!\perp$ 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

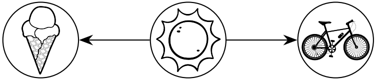
  (a) ♏ $\qquad\qquad$ (b) 🎓 $\longrightarrow$ \$

  Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ \$ (zależne)

- **Trzy zmienne:**

  (a) 

  Mamy $\not\perp\!\!\!\perp$, ale $\perp\!\!\!\perp \mid$ (warunkowa niezależność)

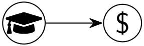
  (b) 

  Tutaj $\not\perp\!\!\!\perp$, ale $\perp\!\!\!\perp \mid$

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a) 

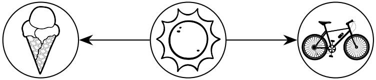
(b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz  $\not\perp\!\!\!\perp$ \$ (zależne)

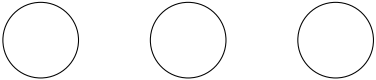
- **Trzy zmienne:**

(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  |  (warunkowa niezależność)

(b) 

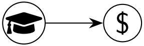
Tutaj  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  | 

(c) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a) 

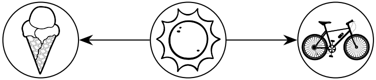
(b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz $\not\perp\!\!\!\perp$ \$ (zależne)

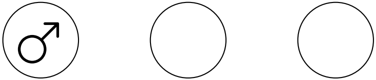
- **Trzy zmienne:**

(a) 

Mamy $\not\perp\!\!\!\perp$, ale $\perp\!\!\!\perp$ $|$ (warunkowa niezależność)

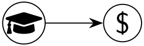
(b) 

Tutaj $\not\perp\!\!\!\perp$, ale $\perp\!\!\!\perp$ $|$

(c) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

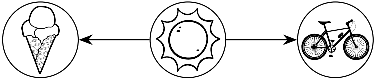
(a)  (b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ \$ (zależne)

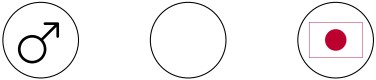
- **Trzy zmienne:**

(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$  (warunkowa niezależność)

(b) 

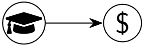
Tutaj  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$ 

(c) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a) 

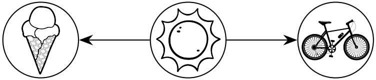
(b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ \$ (zależne)

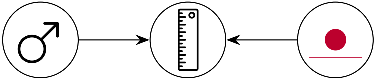
- **Trzy zmienne:**

(a) 

Mamy $\not\perp\!\!\!\perp$, ale $\perp\!\!\!\perp$ | (warunkowa niezależność)

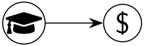
(b) 

Tutaj $\not\perp\!\!\!\perp$, ale $\perp\!\!\!\perp$ |

(c) 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

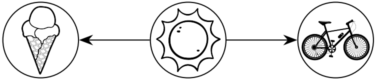
(a) 
(b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz 🎓 $\not\perp\!\!\!\perp$ \$ (zależne)

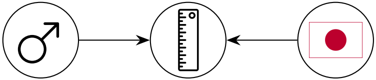
- **Trzy zmienne:**

(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  |  (warunkowa niezależność)

(b) 

Tutaj  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  | 

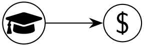
(c) 

Tutaj ♂ $\perp\!\!\!\perp$ 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a) 

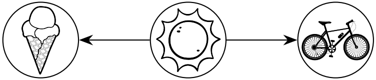
(b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz  $\not\perp\!\!\!\perp$ \$ (zależne)

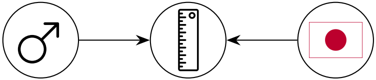
- **Trzy zmienne:**

(a) 

Mamy  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$  (warunkowa niezależność)

(b) 

Tutaj  $\not\perp\!\!\!\perp$ , ale  $\perp\!\!\!\perp$  $|$ 

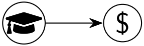
(c) 

Tutaj ♂ $\perp\!\!\!\perp$ , ale ♂ $\perp\!\!\!\perp$  $|$ 

Przykłady małych sieci bayesowskich

- **Dwie zmienne:**

(a) 

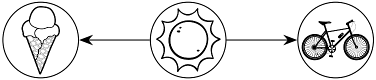
(b) 

Piszemy ♏ $\perp\!\!\!\perp$ \$ (niezależne) oraz $\not\perp\!\!\!\perp$ \$ (zależne)

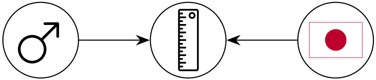
- **Trzy zmienne:**

(a)  **[łańcuch]**

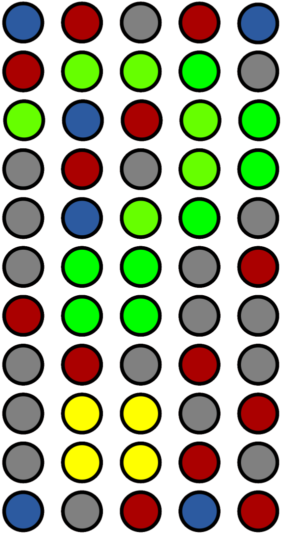
Mamy $\not\perp\!\!\!\perp$, ale $\perp\!\!\!\perp$ | (warunkowa niezależność)

(b)  **[widelec]**

Tutaj $\not\perp\!\!\!\perp$, ale $\perp\!\!\!\perp$ |

(c)  **[zderzacz]**

Tutaj ♂ $\perp\!\!\!\perp$, ale ♂ $\perp\!\!\!\perp$ |



Pytanie

Jak wygląda przepływ informacji w sieci bayesowskiej?

Pytanie

Jak wygląda przepływ informacji w sieci bayesowskiej?

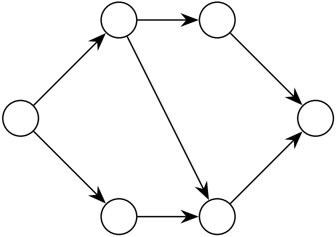
(tzn. jakie warunkowe niezależności są przez nią implikowane?)

Pytanie

Jak wygląda przepływ informacji w sieci bayesowskiej?

(tzn. jakie warunkowe niezależności są przez nią implikowane?)

Przykład:

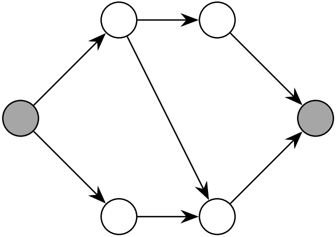


Pytanie

Jak wygląda przepływ informacji w sieci bayesowskiej?

(tzn. jakie warunkowe niezależności są przez nią implikowane?)

Przykład:

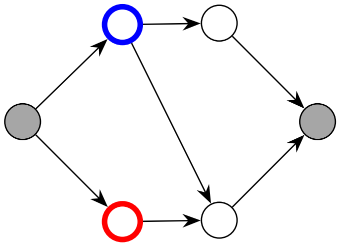


Pytanie

Jak wygląda przepływ informacji w sieci bayesowskiej?

(tzn. jakie warunkowe niezależności są przez nią implikowane?)

Przykład:



Czy ○ $\perp\!\!\!\perp$ ○ | ●?

Przepływ informacji w sieci bayesowskiej

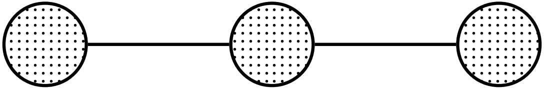
- Informacja może płynąć po krawędziach sieci w obu kierunkach

Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

Legenda:

: warunkowe : niewarunkowe : brak informacji

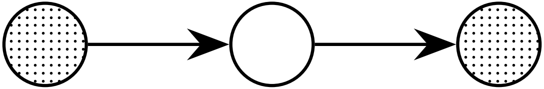


Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

Legenda:

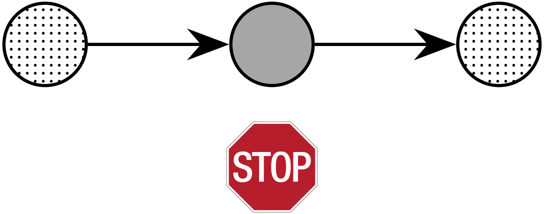
: warunkowe : niewarunkowe : brak informacji



Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

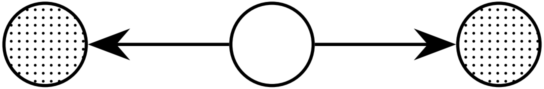
Legenda:
: warunkowe : niewarunkowe : brak informacji



Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

Legenda:
: warunkowe : niewarunkowe : brak informacji

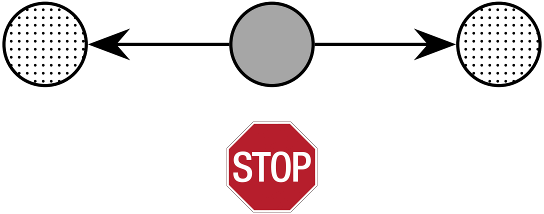


Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

Legenda:

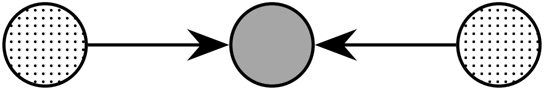
: warunkowe   : niewarunkowe   : brak informacji



Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

Legenda:
: warunkowe : niewarunkowe : brak informacji

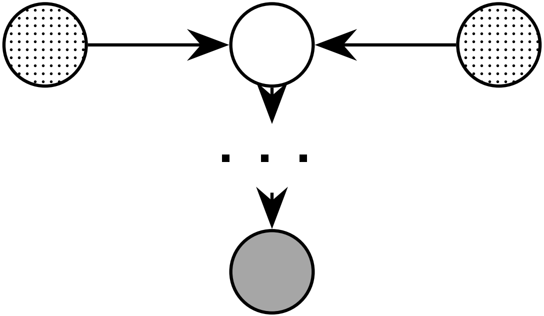


Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

Legenda:

: warunkowe : niewarunkowe : brak informacji

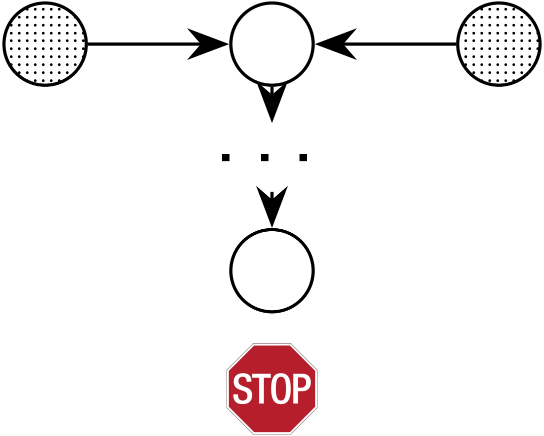


Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

Legenda:

: warunkowe   : niewarunkowe   : brak informacji

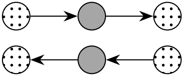
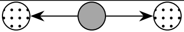
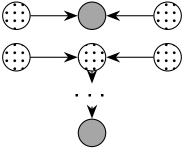


Przepływ informacji w sieci bayesowskiej

- Informacja może płynąć po krawędziach sieci w obu kierunkach

Legenda:

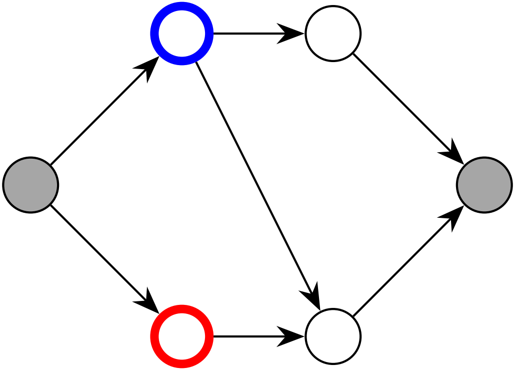
: warunkowe : brak informacji

		
łańcuch		w p. przyp.
widelec		
zderzacz	w p. przyp.	

Przykład

Legenda:

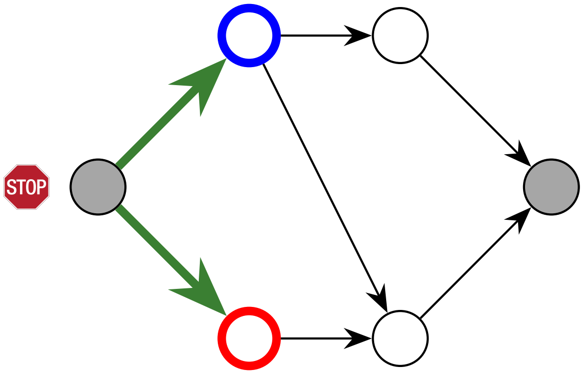
: zmienne „zwarunkowane”   : pozostałe



Przykład

Legenda:

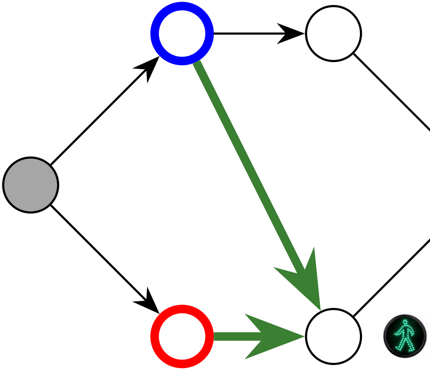
: zmienne „zwarunkowane” : pozostałe



Przykład

Legenda:

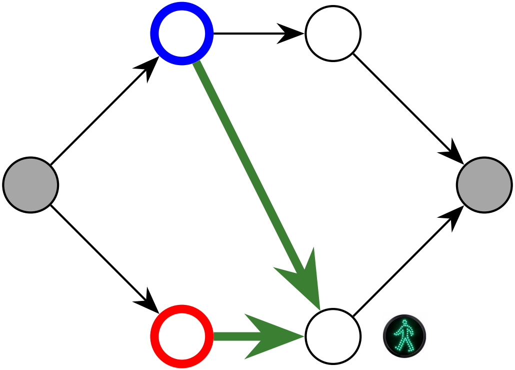
●: zmienne „zwarunkowane” ○: pozostałe



Przykład

Legenda:

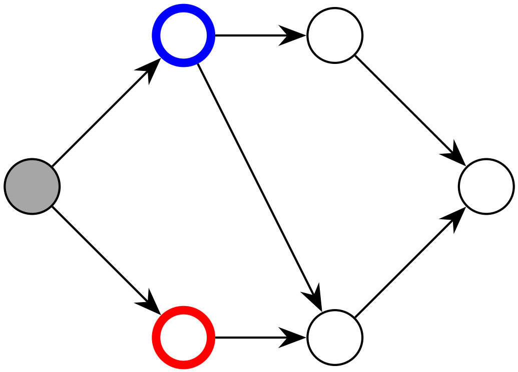
: zmienne „zwarunkowane”   : pozostałe



Przykład

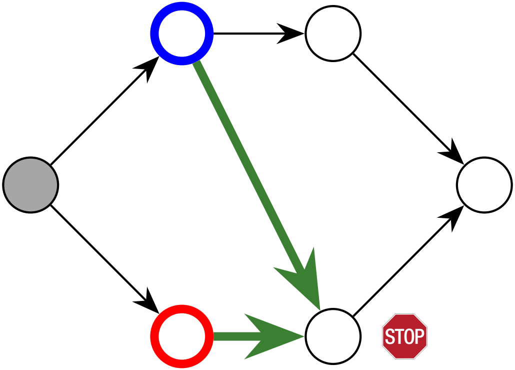
Legenda:

●: zmienne „zwarunkowane” ○: pozostałe



Przykład

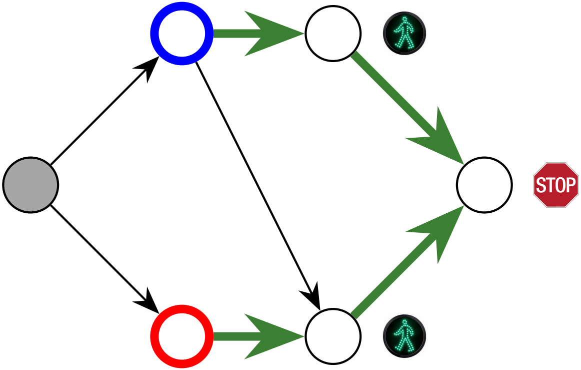
Legenda:
: zmienne „zwarunkowane” : pozostałe



Przykład

Legenda:

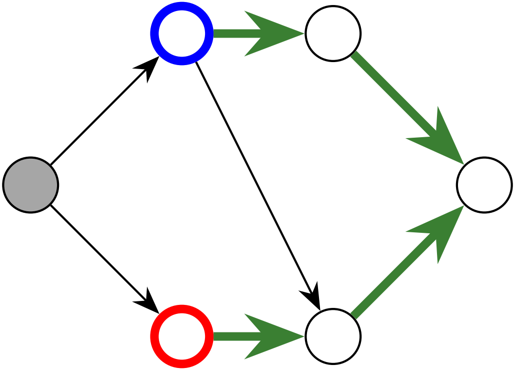
⬤: zmienne „zwarunkowane” ◯: pozostałe

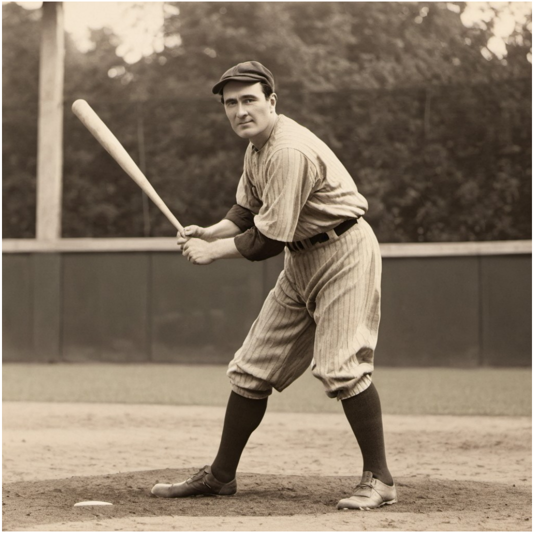


Przykład

Legenda:

: zmienne „zwarunkowane”   : pozostałe





Bayesball :)

Kilka komentarzy

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych

Kilka komentarzy

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych
- brak opisanego wcześniej przepływu informacji nosi nazwę d***-separacji*** (ozn. $A \not\rightarrow B \mid S$)

Kilka komentarzy

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych
- brak opisanego wcześniej przepływu informacji nosi nazwę ***d-separacji*** (ozn. $A \not\to B \mid S$)
- **Twierdzenie Pearla:** d-separacja implikuje warunkową niezależność

Kilka komentarzy

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych
- brak opisanego wcześniej przepływu informacji nosi nazwę ***d-separacji*** (ozn. $A \not\to B \mid S$)
- **Twierdzenie Pearla:** d-separacja implikuje warunkową niezależność
- dalsza część wykładu poświęcona będzie formalizacji twierdzenia Pearla oraz jego dowodowi

Kilka komentarzy

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych
- brak opisanego wcześniej przepływu informacji nosi nazwę ***d-separacji*** (ozn. $A \not\to B \mid S$)
- **Twierdzenie Pearla:** d-separacja implikuje warunkową niezależność
- dalsza część wykładu poświęcona będzie formalizacji twierdzenia Pearla oraz jego dowodowi
- schemat dowodu

d-separacja w grafie skierowanym $\longrightarrow$ warunkowa niezależność

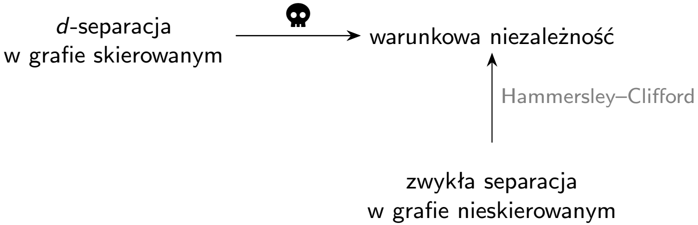
Kilka komentarzy

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych
- brak opisanego wcześniej przepływu informacji nosi nazwę ***d-separacji*** (ozn. $A \not\to B \mid S$)
- **Twierdzenie Pearla:** d-separacja implikuje warunkową niezależność
- dalsza część wykładu poświęcona będzie formalizacji twierdzenia Pearla oraz jego dowodowi
- schemat dowodu

d-separacja w grafie skierowanym ☠⟶ warunkowa niezależność

Kilka komentarzy

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych
- brak opisanego wcześniej przepływu informacji nosi nazwę ***d-separacji*** (ozn. $A \not\to B \mid S$)
- **Twierdzenie Pearla:** d-separacja implikuje warunkową niezależność
- dalsza część wykładu poświęcona będzie formalizacji twierdzenia Pearla oraz jego dowodowi
- schemat dowodu



Kilka komentarzy

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych
- brak opisanego wcześniej przepływu informacji nosi nazwę ***d-separacji*** (ozn. $A \not\to B \mid S$)
- **Twierdzenie Pearla:** d-separacja implikuje warunkową niezależność
- dalsza część wykładu poświęcona będzie formalizacji twierdzenia Pearla oraz jego dowodowi
- schemat dowodu

$$\begin{tikzcd}
\text{d-separacja w grafie skierowanym} \arrow[r] \arrow[d, leftrightarrow] & \text{warunkowa niezależność} \\
\text{pewna zwykła separacja} & \text{zwykła separacja w grafie nieskierowanym} \arrow[u, "\text{Hammersley–Clifford}"']
\end{tikzcd}$$

- termin ***sieci bayesowskie*** został ukuty przez Judeę Pearla (laureat Nagrody Turinga 2011) w latach 80-tych
- brak opisanego wcześniej przepływu informacji nosi nazwę ***d*-separacji** (ozn. $A \not\rightarrow B \mid S$)
- **Twierdzenie Pearla:** d-separacja implikuje warunkową niezależność
- dalsza część wykładu poświęcona będzie formalizacji twierdzenia Pearla oraz jego dowodowi
- schemat dowodu

$$
\begin{tikzcd}
\text{d-separacja w grafie skierowanym} \arrow[r] \arrow[d, leftrightarrow] & \text{warunkowa niezależność} \\
\text{pewna zwykła separacja} \arrow[r, dash, dashed] & \text{zwykła separacja w grafie nieskierowanym} \arrow[u, "\text{Hammersley–Clifford}"']
\end{tikzcd}
$$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X

(Nie)zależności zmiennych + notacja

- $X, Y, Z, ...$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \ldots = p_X(6) = \frac{1}{6}$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \ldots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x, y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \dots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \dots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x, y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

-
  - $p_X(x) = \sum_y p_{X,Y}(x, y)$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \dots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \dots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x,y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y)$$

-
  - $p_X(x) = \sum_y p_{X,Y}(x,y)$, $p_Y(y) = \sum_x p_{X,Y}(x,y)$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \ldots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x, y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

-
  - $p_X(x) = \sum_y p_{X,Y}(x, y)$, $p_Y(y) = \sum_x p_{X,Y}(x, y)$
  - $p_{X\,|\,Y}(x\,|\,y) := \frac{p_{X,Y}(x,y)}{p_Y(y)}$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \ldots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x, y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

-
    - $p_X(x) = \sum_y p_{X,Y}(x, y)$, $p_Y(y) = \sum_x p_{X,Y}(x, y)$
    - $p_{X\,|\,Y}(x\,|\,y) := \frac{p_{X,Y}(x,y)}{p_Y(y)}$ (szansa na $X = x$ pod warunkiem $Y = y$)

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \ldots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x, y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

-
  - $p_X(x) = \sum_y p_{X,Y}(x, y)$, $p_Y(y) = \sum_x p_{X,Y}(x, y)$
  - $p_{X \mid Y}(x \mid y) := \frac{p_{X,Y}(x,y)}{p_Y(y)}$ (szansa na $X = x$ pod warunkiem $Y = y$)
- ***niezależność*** X i Y to stwierdzenie $p_{X \mid Y}(x \mid y) = p_X(x)$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \ldots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x, y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

-
  - $p_X(x) = \sum_y p_{X,Y}(x, y)$, $p_Y(y) = \sum_x p_{X,Y}(x, y)$
  - $p_{X\,|\,Y}(x\,|\,y) := \frac{p_{X,Y}(x,y)}{p_Y(y)}$ (szansa na $X = x$ pod warunkiem $Y = y$)
- ***niezależność*** X i Y to stwierdzenie $p_{X\,|\,Y}(x\,|\,y) = p_X(x)$, czyli

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$$

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \ldots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x, y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

-
  - $p_X(x) = \sum_y p_{X,Y}(x, y)$, $p_Y(y) = \sum_x p_{X,Y}(x, y)$
  - $p_{X\,|\,Y}(x\,|\,y) := \frac{p_{X,Y}(x,y)}{p_Y(y)}$ (szansa na $X = x$ pod warunkiem $Y = y$)
- ***niezależność*** X i Y to stwierdzenie $p_{X\,|\,Y}(x\,|\,y) = p_X(x)$, czyli

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$$

  Równoważnie

$$p_{X,Y}(x, y) = \phi(x) \cdot \psi(y)$$

  dla pewnych funkcji ϕ, ψ.

(Nie)zależności zmiennych + notacja

- $X, Y, Z, \ldots$ – zmienne losowe (dyskretne)
- $p_X(x)$ – ***funkcja prawdopodobieństwa*** zmiennej X, $p_X(x) = \mathbb{P}(X = x)$
- *Przykład.* X – wynik rzutu kością sześcienną
  $p_X(1) = p_X(2) = \ldots = p_X(6) = \frac{1}{6}$
- $p_{X,Y}(x, y)$ – *łączna* funkcja prawdopodobieństwa dla zmiennych X, Y

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

-
  - $p_X(x) = \sum_y p_{X,Y}(x, y)$, $p_Y(y) = \sum_x p_{X,Y}(x, y)$
  - $p_{X\,|\,Y}(x\,|\,y) := \frac{p_{X,Y}(x,y)}{p_Y(y)}$ (szansa na $X = x$ pod warunkiem $Y = y$)
- ***niezależność*** X i Y to stwierdzenie $p_{X\,|\,Y}(x\,|\,y) = p_X(x)$, czyli

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$$

  Równoważnie

$$p_{X,Y}(x, y) = \phi(x) \cdot \psi(y)$$

  dla pewnych funkcji ϕ, ψ. Piszemy $X \perp\!\!\!\perp Y$.

(Nie)zależności zmiennych + notacja

- ***niezależność*** X i Y to stwierdzenie $p_{X \mid Y}(x \mid y) = p_X(x)$, czyli

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$$

Równoważnie

$$p_{X,Y}(x, y) = \phi(x) \cdot \psi(y)$$

dla pewnych funkcji ϕ, ψ. Piszemy $X \perp\!\!\!\perp Y$.

(Nie)zależności zmiennych + notacja

- ***niezależność*** X i Y to stwierdzenie $p_{X\,|\,Y}(x\,|\,y) = p_X(x)$, czyli

  $$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$$

  Równoważnie

  $$p_{X,Y}(x,y) = \phi(x) \cdot \psi(y)$$

  dla pewnych funkcji ϕ, ψ. Piszemy $X \perp\!\!\!\perp Y$.

- ***niezależność X i Y pod warunkiem Z*** to stwierdzenie $p_{X\,|\,Y,Z}(x\,|\,y,z) = p_{X\,|\,Z}(x\,|\,z)$

(Nie)zależności zmiennych + notacja

- ***niezależność*** X i Y to stwierdzenie $p_{X\,|\,Y}(x\,|\,y) = p_X(x)$, czyli

  $$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$$

  Równoważnie

  $$p_{X,Y}(x,y) = \phi(x) \cdot \psi(y)$$

  dla pewnych funkcji ϕ, ψ. Piszemy $X \perp\!\!\!\perp Y$.

- ***niezależność X i Y pod warunkiem Z*** to stwierdzenie $p_{X\,|\,Y,Z}(x\,|\,y,z) = p_{X\,|\,Z}(x\,|\,z)$, czyli

  $$p_{X,Y,Z}(x,y,z) = p_{X,Z}(x,z) \cdot p_{Y,Z}(y,z) / p_Z(z)$$

(Nie)zależności zmiennych + notacja

- ***niezależność*** X i Y to stwierdzenie $p_{X \mid Y}(x \mid y) = p_X(x)$, czyli

  $$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$$

  Równoważnie

  $$p_{X,Y}(x, y) = \phi(x) \cdot \psi(y)$$

  dla pewnych funkcji ϕ, ψ. Piszemy $X \perp\!\!\!\perp Y$.

- ***niezależność X i Y pod warunkiem Z*** to stwierdzenie $p_{X \mid Y,Z}(x \mid y, z) = p_{X \mid Z}(x \mid z)$, czyli

  $$p_{X,Y,Z}(x, y, z) = p_{X,Z}(x, z) \cdot p_{Y,Z}(y, z)/p_Z(z)$$

  Równoważnie

  $$p_{X,Y,Z}(x, y, z) = \Phi(x, z) \cdot \Psi(y, z)$$

  dla pewnych funkcji Φ, Ψ.

(Nie)zależności zmiennych + notacja

- ***niezależność*** X i Y to stwierdzenie $p_{X\,|\,Y}(x\,|\,y) = p_X(x)$, czyli

  $$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$$

  Równoważnie

  $$p_{X,Y}(x,y) = \phi(x) \cdot \psi(y)$$

  dla pewnych funkcji ϕ, ψ. Piszemy $X \perp\!\!\!\perp Y$.

- ***niezależność X i Y pod warunkiem Z*** to stwierdzenie $p_{X\,|\,Y,Z}(x\,|\,y,z) = p_{X\,|\,Z}(x\,|\,z)$, czyli

  $$p_{X,Y,Z}(x,y,z) = p_{X,Z}(x,z) \cdot p_{Y,Z}(y,z) / p_Z(z)$$

  Równoważnie

  $$p_{X,Y,Z}(x,y,z) = \Phi(x,z) \cdot \Psi(y,z)$$

  dla pewnych funkcji Φ, Ψ. Piszemy $X \perp\!\!\!\perp Y\,|\,Z$.

(Nie)zależności zmiennych + notacja

- ***niezależność*** X i Y to stwierdzenie $p_{X\,|\,Y}(x\,|\,y) = p_X(x)$, czyli

  $$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$$

  Równoważnie

  $$p_{X,Y}(x,y) = \phi(x) \cdot \psi(y)$$

  dla pewnych funkcji ϕ, ψ. Piszemy $X \perp\!\!\!\perp Y$.

- ***niezależność X i Y pod warunkiem Z*** to stwierdzenie $p_{X\,|\,Y,Z}(x\,|\,y,z) = p_{X\,|\,Z}(x\,|\,z)$, czyli

  $$p_{X,Y,Z}(x,y,z) = p_{X,Z}(x,z) \cdot p_{Y,Z}(y,z) / p_Z(z)$$

  Równoważnie

  $$p_{X,Y,Z}(x,y,z) = \Phi(x,z) \cdot \Psi(y,z)$$

  dla pewnych funkcji Φ, Ψ. Piszemy $X \perp\!\!\!\perp Y\,|\,Z$.

- pomijamy dolny indeks funkcji p

(Nie)zależności zmiennych + notacja

- ***niezależność*** $\boldsymbol{X}$ i $\boldsymbol{Y}$ to stwierdzenie $p_{X \mid Y}(\boldsymbol{x} \mid \boldsymbol{y}) = p_X(\boldsymbol{x})$, czyli

$$p_{X,Y}(\boldsymbol{x}, \boldsymbol{y}) = p_X(\boldsymbol{x}) \cdot p_Y(\boldsymbol{y})$$

  Równoważnie

$$p_{X,Y}(\boldsymbol{x}, \boldsymbol{y}) = \phi(\boldsymbol{x}) \cdot \psi(\boldsymbol{y})$$

  dla pewnych funkcji ϕ, ψ. Piszemy $\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}$.

- ***niezależność X i Y pod warunkiem Z*** to stwierdzenie $p_{X \mid Y,Z}(\boldsymbol{x} \mid \boldsymbol{y}, \boldsymbol{z}) = p_{X \mid Z}(\boldsymbol{x} \mid \boldsymbol{z})$, czyli

$$p_{X,Y,Z}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = p_{X,Z}(\boldsymbol{x}, \boldsymbol{z}) \cdot p_{Y,Z}(\boldsymbol{y}, \boldsymbol{z}) / p_Z(\boldsymbol{z})$$

  Równoważnie

$$p_{X,Y,Z}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \Phi(\boldsymbol{x}, \boldsymbol{z}) \cdot \Psi(\boldsymbol{y}, \boldsymbol{z})$$

  dla pewnych funkcji Φ, Ψ. Piszemy $\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}$.

- pomijamy dolny indeks funkcji p
- to samo zachodzi dla wektorów losowych $\boldsymbol{X}$, $\boldsymbol{Y}$ i $\boldsymbol{Z}$

Notacji ciąg dalszy

- Niech $\boldsymbol{x} = (x_1, \ldots, x_n)$, wtedy $\boldsymbol{x}_A = (x_i \colon i \in A)$, np.

$$\boldsymbol{x}_{\{2,5,7\}} := (x_2, x_5, x_7)$$

Notacji ciąg dalszy

- Niech $\boldsymbol{x} = (x_1, \ldots, x_n)$, wtedy $\boldsymbol{x}_A = (x_i \colon i \in A)$, np.

$$\boldsymbol{x}_{\{2,5,7\}} := (x_2, x_5, x_7)$$

Ta sama konwencja dotyczy ciągu zmiennych losowych.

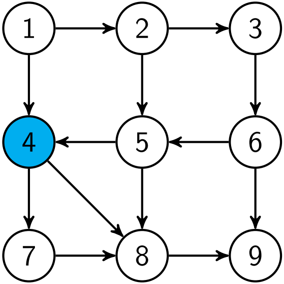
Notacji ciąg dalszy

- Niech $\boldsymbol{x} = (x_1, \ldots, x_n)$, wtedy $\boldsymbol{x}_A = (x_i \colon i \in A)$, np.

$$\boldsymbol{x}_{\{2,5,7\}} := (x_2, x_5, x_7)$$

  Ta sama konwencja dotyczy ciągu zmiennych losowych.

- Definicje grafowe:



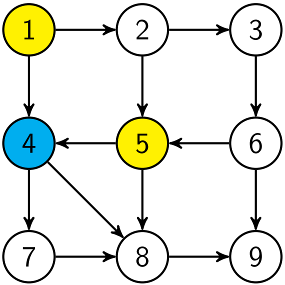
Notacji ciąg dalszy

- Niech $\boldsymbol{x} = (x_1, \ldots, x_n)$, wtedy $\boldsymbol{x}_A = (x_i \colon i \in A)$, np.

$$\boldsymbol{x}_{\{2,5,7\}} := (x_2, x_5, x_7)$$

  Ta sama konwencja dotyczy ciągu zmiennych losowych.

- Definicje grafowe:



  - rodzice
    $\mathrm{pa}(4) = \{1, 5\}$

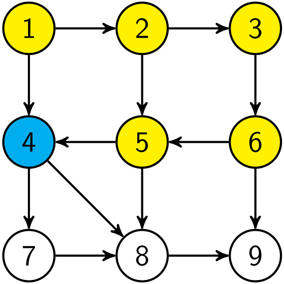
Notacji ciąg dalszy

- Niech $\boldsymbol{x} = (x_1, \ldots, x_n)$, wtedy $\boldsymbol{x}_A = (x_i \colon i \in A)$, np.

$$\boldsymbol{x}_{\{2,5,7\}} := (x_2, x_5, x_7)$$

  Ta sama konwencja dotyczy ciągu zmiennych losowych.

- Definicje grafowe:



  - rodzice
    $\mathrm{pa}(4) = \{1, 5\}$
  - przodkowie
    $\mathrm{an}(4) = \{1, 2, 3, 5, 6\}$

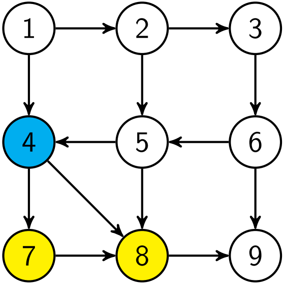
Notacji ciąg dalszy

- Niech $\boldsymbol{x} = (x_1, \ldots, x_n)$, wtedy $\boldsymbol{x}_A = (x_i \colon i \in A)$, np.

$$\boldsymbol{x}_{\{2,5,7\}} := (x_2, x_5, x_7)$$

  Ta sama konwencja dotyczy ciągu zmiennych losowych.

- Definicje grafowe:



  - rodzice
    $\mathrm{pa}(4) = \{1, 5\}$
  - przodkowie
    $\mathrm{an}(4) = \{1, 2, 3, 5, 6\}$
  - dzieci
    $\mathrm{ch}(4) = \{7, 8\}$

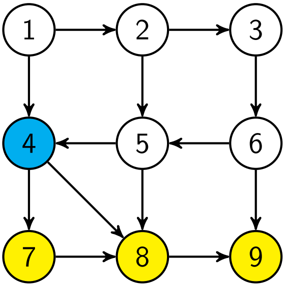
Notacji ciąg dalszy

- Niech $\boldsymbol{x} = (x_1, \ldots, x_n)$, wtedy $\boldsymbol{x}_A = (x_i : i \in A)$, np.

$$\boldsymbol{x}_{\{2,5,7\}} := (x_2, x_5, x_7)$$

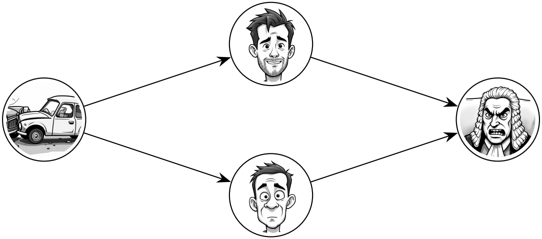
  Ta sama konwencja dotyczy ciągu zmiennych losowych.

- Definicje grafowe:



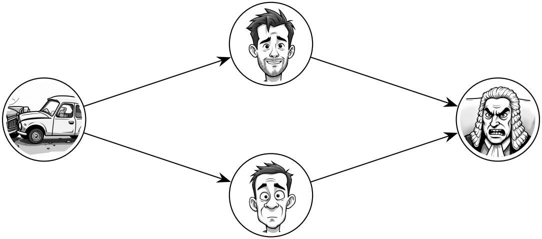
  - rodzice
    $\mathrm{pa}(4) = \{1, 5\}$
  - przodkowie
    $\mathrm{an}(4) = \{1, 2, 3, 5, 6\}$
  - dzieci
    $\mathrm{ch}(4) = \{7, 8\}$
  - potomkowie
    $\mathrm{de}(4) = \{7, 8, 9\}$

Co to jest sieć bayesowska?



- jeśli znane są zeznania, prawda jest nieistotna dla werdyktu

Co to jest sieć bayesowska?

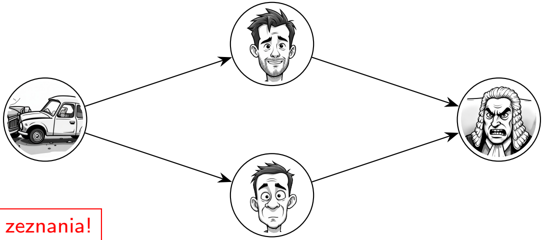


- jeśli znane są zeznania, prawda jest nieistotna dla werdyktu

Definicja. Rozważmy zmienne $X_1, \ldots, X_n$ o funkcji prawdopodobieństwa p. Powiemy, że acykliczny graf skierowany $\mathcal{G}$ o wierzchołkach w $1, \ldots, n$ jest *siecią bayesowską* dla p, jeśli $\mathrm{pa}(i) \subseteq \{1, \ldots, i-1\}$ oraz

$$X_i \perp\!\!\!\perp \boldsymbol{X}_{\{1,\ldots,i-1\} \setminus \mathrm{pa}(i)} \,\big|\, \boldsymbol{X}_{\mathrm{pa}(i)},$$

Co to jest sieć bayesowska?



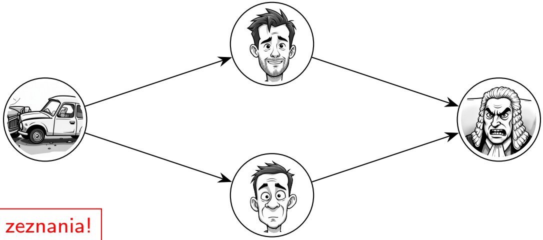
Oba zeznania!

- jeśli znane są zeznania, prawda jest nieistotna dla werdyktu

Definicja. Rozważmy zmienne $X_1, \dots, X_n$ o funkcji prawdopodobieństwa p. Powiemy, że acykliczny graf skierowany $\mathcal{G}$ o wierzchołkach w $1, \dots, n$ jest *siecią bayesowską* dla p, jeśli $\mathrm{pa}(i) \subseteq \{1, \dots, i-1\}$ oraz

$$X_i \perp\!\!\!\perp \boldsymbol{X}_{\{1,\dots,i-1\} \setminus \mathrm{pa}(i)} \,\big|\, \boldsymbol{X}_{\mathrm{pa}(i)},$$

Co to jest sieć bayesowska?



Oba zeznania!

- jeśli znane są zeznania, prawda jest nieistotna dla werdyktu

Definicja. Rozważmy zmienne $X_1, \ldots, X_n$ o funkcji prawdopodobieństwa p. Powiemy, że acykliczny graf skierowany $\mathcal{G}$ o wierzchołkach w $1, \ldots, n$ jest *siecią bayesowską* dla p, jeśli $\mathrm{pa}(i) \subseteq \{1, \ldots, i-1\}$ oraz

$$X_i \perp\!\!\!\perp \boldsymbol{X}_{\{1,\ldots,i-1\} \setminus \mathrm{pa}(i)} \,\big|\, \boldsymbol{X}_{\mathrm{pa}(i)},$$

a także $X_i \not\perp\!\!\!\perp \boldsymbol{X}_{\{1,\ldots,i-1\} \setminus A} \,\big|\, X_A$ dla dowolnego $A \subsetneq \mathrm{pa}(A)$

Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe

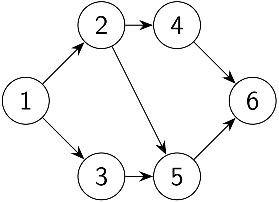
Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe

$$\begin{aligned} p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\ p(x_2 \mid x_1)\cdot \\ p(x_3 \mid x_1, x_2)\cdot \\ p(x_4 \mid x_1, x_2, x_3)\cdot \\ p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\ p(x_6 \mid x_1, x_2, x_3, x_4, x_5) \end{aligned}$$

Faktoryzacja funkcji prawdopodobieństwa

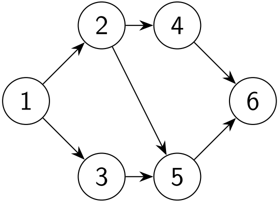
$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe



$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5)
\end{aligned}
$$

Faktoryzacja funkcji prawdopodobieństwa

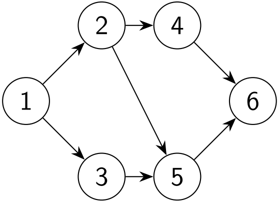
$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe



$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
\boxed{X_3 \perp\!\!\!\perp X_2 \mid X_1} \longrightarrow p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5)
\end{aligned}
$$

Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe

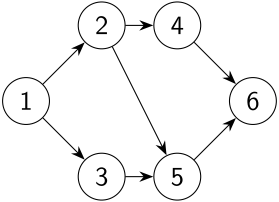


$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = {} & p(x_1)\cdot \\
& p(x_2 \mid x_1)\cdot \\
& p(x_3 \mid x_1, x_2)\cdot \\
& p(x_4 \mid x_1, x_2, x_3)\cdot \\
& p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
& p(x_6 \mid x_1, x_2, x_3, x_4, x_5)
\end{aligned}
$$

$\boxed{X_3 \perp\!\!\!\perp X_2 \mid X_1}$

Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe

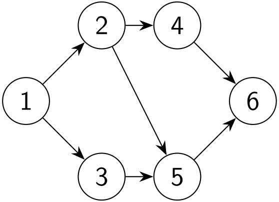


$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5)
\end{aligned}
$$

$\boxed{X_4 \perp\!\!\!\perp (X_1, X_3) \mid X_2}$

Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe

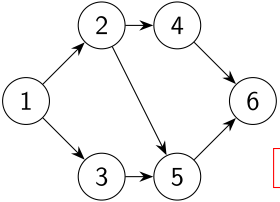


$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5)
\end{aligned}
$$

$\boxed{X_4 \perp\!\!\!\perp (X_1, X_3) \mid X_2}$

Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe

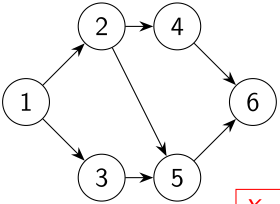


$$\begin{aligned} p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot & \\ p(x_2 \mid x_1)\cdot & \\ p(x_3 \mid x_1, x_2)\cdot & \\ p(x_4 \mid x_1, x_2, x_3)\cdot & \\ p(x_5 \mid x_1, x_2, x_3, x_4)\cdot & \\ p(x_6 \mid x_1, x_2, x_3, x_4, x_5) & \end{aligned}$$

$$\boxed{X_5 \perp\!\!\!\perp (X_1, X_4) \mid (X_2, X_3)}$$

Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe

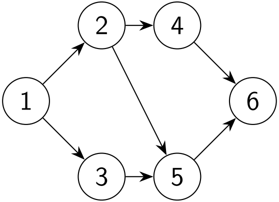


$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5)
\end{aligned}
$$

$$\boxed{X_6 \perp\!\!\!\perp (X_1, X_2, X_3) \mid (X_4, X_5)} \longrightarrow p(x_6 \mid x_1, x_2, x_3, x_4, x_5)$$

Faktoryzacja funkcji prawdopodobieństwa

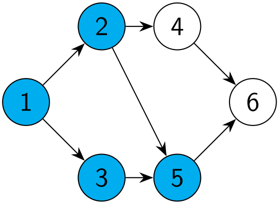
$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe



$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5) = \\
= \prod_{i=1}^{6} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})
\end{aligned}
$$

Faktoryzacja funkcji prawdopodobieństwa

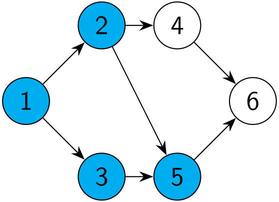
$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe



$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5) = \\
= \prod_{i=1}^{6} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})
\end{aligned}
$$

Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe



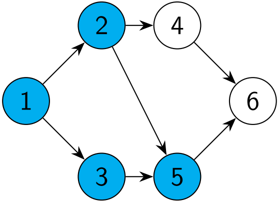
$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5) = \\
= \prod_{i=1}^{6} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})
\end{aligned}
$$

Zauważmy ponadto, że

$$p(x_1, x_2, x_3, x_5) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1)p(x_5 \mid x_3)$$

Faktoryzacja funkcji prawdopodobieństwa

$X_1, X_2, X_3, X_4, X_5, X_6$ – zmienne losowe



$$
\begin{aligned}
p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)\cdot \\
p(x_2 \mid x_1)\cdot \\
p(x_3 \mid x_1, x_2)\cdot \\
p(x_4 \mid x_1, x_2, x_3)\cdot \\
p(x_5 \mid x_1, x_2, x_3, x_4)\cdot \\
p(x_6 \mid x_1, x_2, x_3, x_4, x_5) = \\
= \prod_{i=1}^{6} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})
\end{aligned}
$$

Zauważmy ponadto, że

$$p(x_1, x_2, x_3, x_5) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1)p(x_5 \mid x_3) = \prod_{i \in \{1,2,3,5\}} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Faktoryzacja funkcji prawdopodobieństwa

Obserwacja. Jeśli $\mathcal{G}$ jest siecią bayesowską dla zmiennych $X_1, \ldots, X_n$, to

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)}).$$

Obserwacja. Jeśli $\mathcal{G}$ jest siecią bayesowską dla zmiennych $X_1, \ldots, X_n$, oraz A jest ***zbiorem samorodnym***, tzn. $\mathrm{pa}(A) \subseteq A$, to

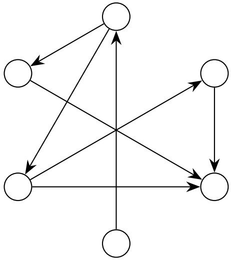
$$p(\boldsymbol{x}_A) = \prod_{i \in A} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)}).$$

Kolejne grafowe definicje

- ***zbiór ancestralny do*** A ***w*** $\mathcal{G}$ to $\mathrm{An}(A) := A \cup \{\mathrm{an}(i)\colon i \in A\}$

Kolejne grafowe definicje

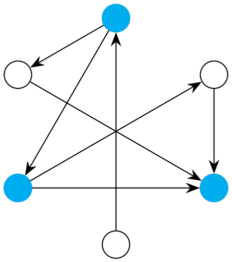
- ***zbiór ancestralny do A w $\mathcal{G}$*** to $\mathrm{An}(A) := A \cup \{\mathrm{an}(i) \colon i \in A\}$
- ***graf indukowany***



$\mathcal{G}$

Kolejne grafowe definicje

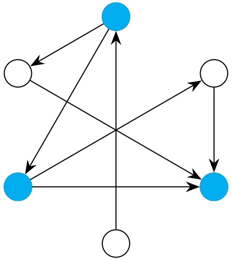
- ***zbiór ancestralny do A w $\mathcal{G}$*** to $\mathrm{An}(A) := A \cup \{\mathrm{an}(i) \colon i \in A\}$
- ***graf indukowany***



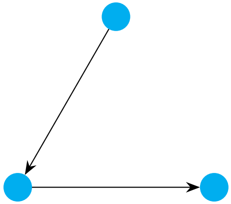
$\mathcal{G}$

Kolejne grafowe definicje

- ***zbiór ancestralny do*** A ***w*** $\mathcal{G}$ to $\mathrm{An}(A) := A \cup \{\mathrm{an}(i) \colon i \in A\}$
- ***graf indukowany***



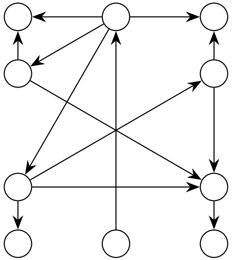
$\longrightarrow$



$\mathcal{G}$ $\qquad$ $\mathcal{G}_A$

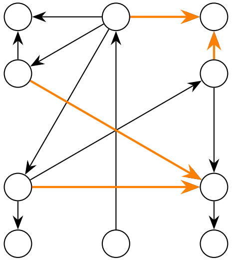
Kolejne grafowe definicje

- ***moralizacja grafu skierowanego***



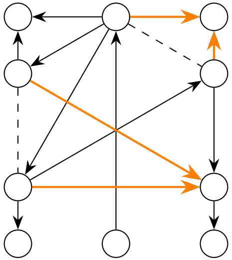
$\mathcal{G}$

- ***moralizacja grafu skierowanego***



$\mathcal{G}$

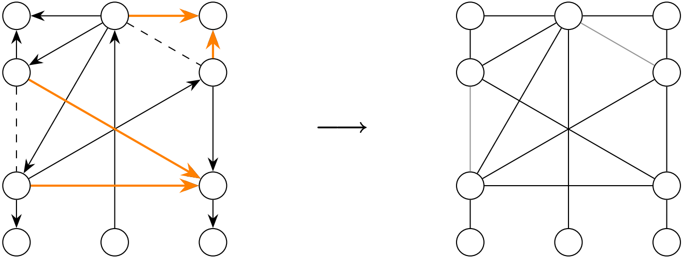
- ***moralizacja grafu skierowanego***



$\mathcal{G}$

Kolejne grafowe definicje

- ***moralizacja grafu skierowanego***

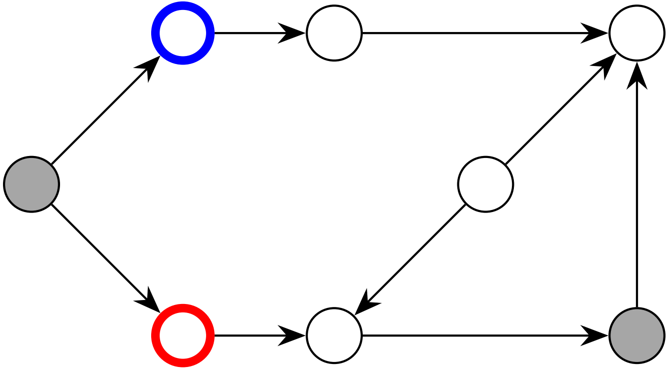


$\mathcal{G}$ $\qquad\qquad$ $\mathcal{G}^{\mathrm{m}}$

Twierdzenie. W grafie skierowanym $\mathcal{G}$ zachodzi $A \not\rightarrow B \mid S$ wtedy i tylko wtedy, gdy w grafie $\mathcal{G}^{\mathrm{m}}_{\mathsf{An}(A \cup B \cup S)}$ każda ścieżka łącząca A i B przecina S (ozn. $A \not\!\!-\!\!\!- B \mid S$).

Twierdzenie. W grafie skierowanym $\mathcal{G}$ zachodzi $A \nrightarrow B \mid S$ wtedy i tylko wtedy, gdy w grafie $\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A \cup B \cup S)}$ każda ścieżka łącząca A i B przecina S (ozn. $A \mathrel{\not\!\!-} B \mid S$).

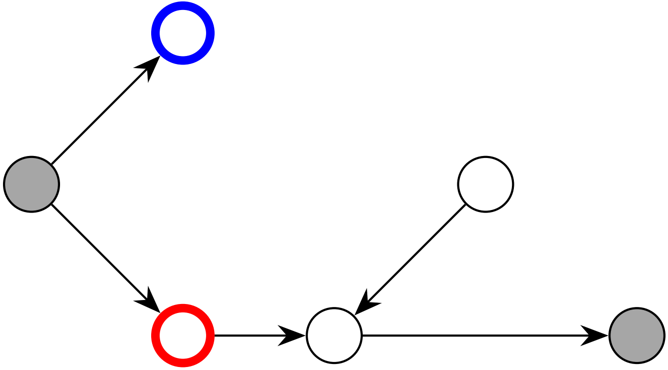
Przykład 1:



$\mathcal{G}$

Twierdzenie. W grafie skierowanym $\mathcal{G}$ zachodzi $A \not\to B \mid S$ wtedy i tylko wtedy, gdy w grafie $\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A \cup B \cup S)}$ każda ścieżka łącząca A i B przecina S (ozn. $A \not\mathrel{-} B \mid S$).

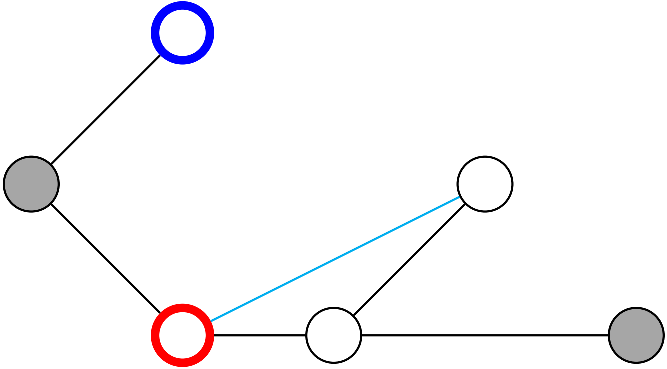
Przykład 1:



$\mathcal{G}_{\mathrm{An}(A \cup B \cup S)}$

Twierdzenie. W grafie skierowanym $\mathcal{G}$ zachodzi $A \not\rightarrow B \mid S$ wtedy i tylko wtedy, gdy w grafie $\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A\cup B\cup S)}$ każda ścieżka łącząca A i B przecina S (ozn. $A \not- B \mid S$).

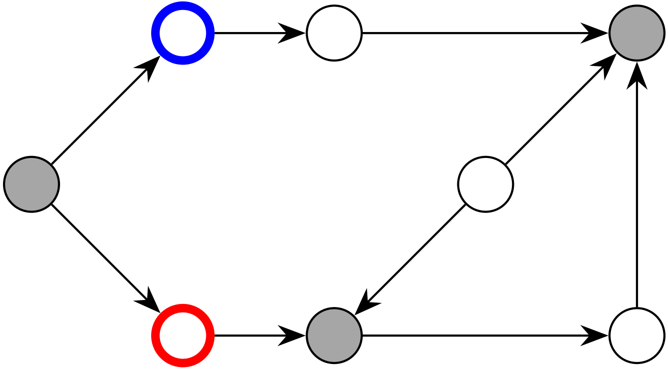
Przykład 1:



$$\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A\cup B\cup S)}$$

Twierdzenie. W grafie skierowanym $\mathcal{G}$ zachodzi $A \not\rightarrow B \mid S$ wtedy i tylko wtedy, gdy w grafie $\mathcal{G}^{\mathrm{m}}_{\mathsf{An}(A \cup B \cup S)}$ każda ścieżka łącząca A i B przecina S (ozn. $A \not\!\!-\!\!- B \mid S$).

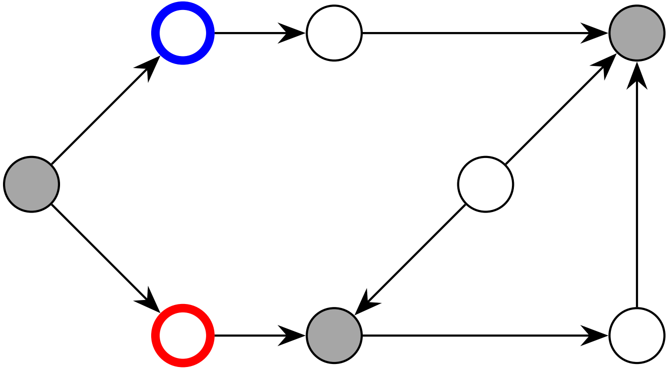
Przykład 2:



$\mathcal{G}$

Twierdzenie. W grafie skierowanym $\mathcal{G}$ zachodzi $A \not\rightarrow B \mid S$ wtedy i tylko wtedy, gdy w grafie $\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A \cup B \cup S)}$ każda ścieżka łącząca A i B przecina S (ozn. $A \not- B \mid S$).

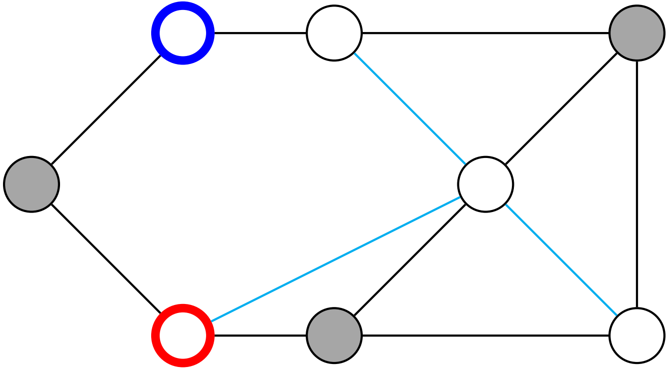
Przykład 2:



$$\mathcal{G}_{\mathrm{An}(A \cup B \cup S)}$$

Twierdzenie. W grafie skierowanym $\mathcal{G}$ zachodzi $A \not\to B \mid S$ wtedy i tylko wtedy, gdy w grafie $\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A \cup B \cup S)}$ każda ścieżka łącząca A i B przecina S (ozn. $A \not\!\!- B \mid S$).

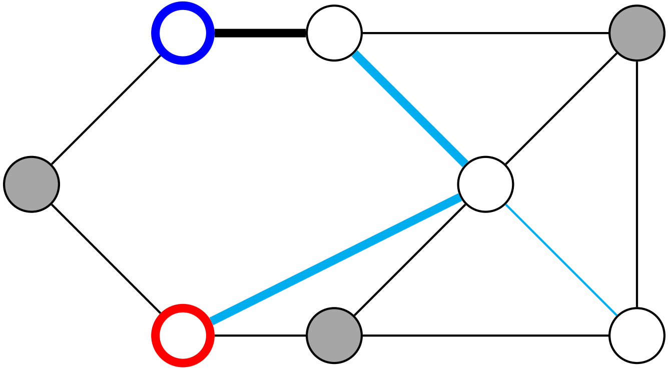
Przykład 2:



$\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A \cup B \cup S)}$

Twierdzenie. W grafie skierowanym $\mathcal{G}$ zachodzi $A \not\rightarrow B \mid S$ wtedy i tylko wtedy, gdy w grafie $\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A \cup B \cup S)}$ każda ścieżka łącząca A i B przecina S (ozn. $A \not\!- B \mid S$).

Przykład 2:



$$\mathcal{G}^{\mathrm{m}}_{\mathrm{An}(A \cup B \cup S)}$$

Twierdzenie. *(Hammersley-Clifford)* Rozważmy zmienne losowe $X_1, \ldots, X_n$ o łącznej funkcji prawdopodobieństwa $p(x_1, \ldots, x_n)$. Niech $\mathcal{H}$ będzie grafem (nieskierowanym) o wierzchołkach $V = 1, ..., n$. Wówczas implikacja

$$A \not\sim B \mid S \Rightarrow \boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid X_S$$

zachodzi dla dowolnych $A, B, S \subseteq V$ wtedy i tylko wtedy, gdy

$$p(\boldsymbol{x}) = \prod_{C \in \mathcal{C}} \xi(\boldsymbol{x}_C) \tag{HC}$$

gdzie $\mathcal{C}$ to rodzina wszystkich klik w grafie $\mathcal{H}$.

Twierdzenie. *(Hammersley-Clifford)* Rozważmy zmienne losowe $X_1, \ldots, X_n$ o łącznej funkcji prawdopodobieństwa $p(x_1, \ldots, x_n)$. Niech $\mathcal{H}$ będzie grafem (nieskierowanym) o wierzchołkach $V = 1, ..., n$. Wówczas implikacja

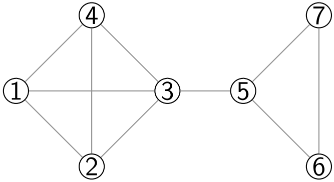
$$A \not\sim B \mid S \Rightarrow \boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid X_S$$

zachodzi dla dowolnych $A, B, S \subseteq V$ wtedy i tylko wtedy, gdy

$$p(\boldsymbol{x}) = \prod_{C \in \mathcal{C}} \xi(\boldsymbol{x}_C) \tag{HC}$$

gdzie $\mathcal{C}$ to rodzina wszystkich klik w grafie $\mathcal{H}$.

Przykład.



Twierdzenie. *(Hammersley-Clifford)* Rozważmy zmienne losowe $X_1, \ldots, X_n$ o łącznej funkcji prawdopodobieństwa $p(x_1, \ldots, x_n)$. Niech $\mathcal{H}$ będzie grafem (nieskierowanym) o wierzchołkach $V = 1, ..., n$. Wówczas implikacja

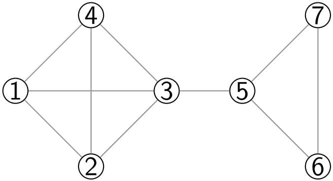
$$A \not\sim B \mid S \Rightarrow \boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid X_S$$

zachodzi dla dowolnych $A, B, S \subseteq V$ wtedy i tylko wtedy, gdy

$$p(\boldsymbol{x}) = \prod_{C \in \mathcal{C}} \xi(\boldsymbol{x}_C) \tag{HC}$$

gdzie $\mathcal{C}$ to rodzina wszystkich klik w grafie $\mathcal{H}$.

Przykład.



$$p(\boldsymbol{x}) = V_a(x_1, x_2, x_3, x_4) \cdot V_b(x_3, x_5) \cdot V_c(x_5, x_6, x_7)$$

Twierdzenie. *(Hammersley-Clifford)* Rozważmy zmienne losowe $X_1, \ldots, X_n$ o łącznej funkcji prawdopodobieństwa $p(x_1, \ldots, x_n)$. Niech $\mathcal{H}$ będzie grafem (nieskierowanym) o wierzchołkach $V = 1, ..., n$. Wówczas implikacja

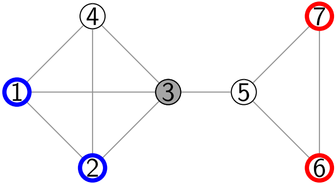
$$A \not\sim B \mid S \Rightarrow \boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid X_S$$

zachodzi dla dowolnych $A, B, S \subseteq V$ wtedy i tylko wtedy, gdy

$$p(\boldsymbol{x}) = \prod_{C \in \mathcal{C}} \xi(\boldsymbol{x}_C) \qquad \text{(HC)}$$

gdzie $\mathcal{C}$ to rodzina wszystkich klik w grafie $\mathcal{H}$.

Przykład.



$$p(\boldsymbol{x}) = V_a(x_1, x_2, x_3, x_4) \cdot V_b(x_3, x_5) \cdot V_c(x_5, x_6, x_7)$$

$$\{X_1, X_2\} \perp\!\!\!\perp \{X_6, X_7\} \mid X_3?$$

Twierdzenie. *(Hammersley-Clifford)* Rozważmy zmienne losowe $X_1, \ldots, X_n$ o łącznej funkcji prawdopodobieństwa $p(x_1, \ldots, x_n)$. Niech $\mathcal{H}$ będzie grafem (nieskierowanym) o wierzchołkach $V = 1, \ldots, n$. Wówczas implikacja

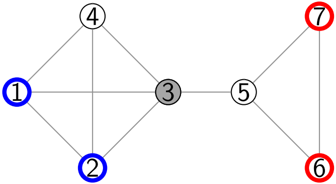
$$A \not\!\!{-}\!\!{-}\, B \mid S \Rightarrow \boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid X_S$$

zachodzi dla dowolnych $A, B, S \subseteq V$ wtedy i tylko wtedy, gdy

$$p(\boldsymbol{x}) = \prod_{C \in \mathcal{C}} \xi(\boldsymbol{x}_C) \tag{HC}$$

gdzie $\mathcal{C}$ to rodzina wszystkich klik w grafie $\mathcal{H}$.

Przykład.



$$p(\boldsymbol{x}) = V_a(x_1, x_2, x_3, x_4) \cdot V_b(x_3, x_5) \cdot V_c(x_5, x_6, x_7)$$

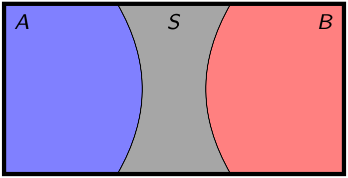
$\{X_1, X_2\} \perp\!\!\!\perp \{X_6, X_7\} \mid X_3$? TAK!

Dowód $\Leftarrow$

Przypadek I. $A \cup B \cup S = V$

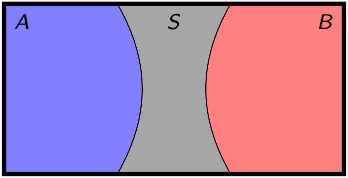
Dowód $\Leftarrow$

Przypadek I. $A \cup B \cup S = V$



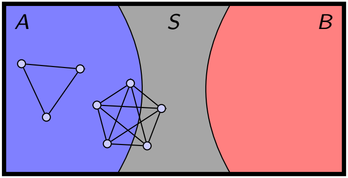
Dowód $\Leftarrow$

Przypadek I. $A \cup B \cup S = V$



$\mathcal{C} =$

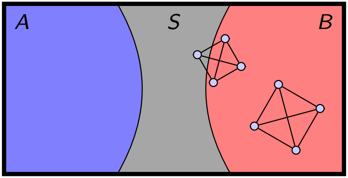
Przypadek I. $A \cup B \cup S = V$



$$\mathcal{C} = \mathcal{C}_A \sqcup$$

Dowód $\Leftarrow$

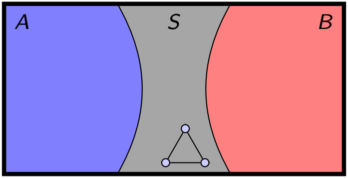
Przypadek I. $A \cup B \cup S = V$



$$\mathcal{C} = \mathcal{C}_A \sqcup \mathcal{C}_B \sqcup$$

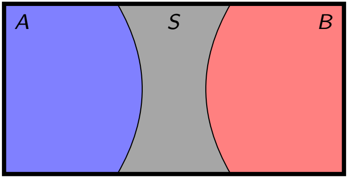
Dowód $\Leftarrow$

Przypadek I. $A \cup B \cup S = V$



$$\mathcal{C} = \mathcal{C}_A \sqcup \mathcal{C}_B \sqcup \mathcal{C}_S$$

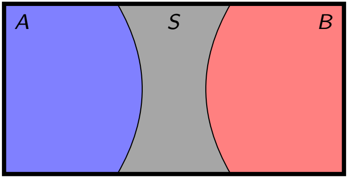
Przypadek I. $A \cup B \cup S = V$



$$\mathcal{C} = \mathcal{C}_A \sqcup \mathcal{C}_B \sqcup \mathcal{C}_S$$

$$\text{(HC)} \quad p(\boldsymbol{x}) = \prod_{C \in \mathcal{C}} V_C(\boldsymbol{x}_C)$$

Przypadek I. $A \cup B \cup S = V$

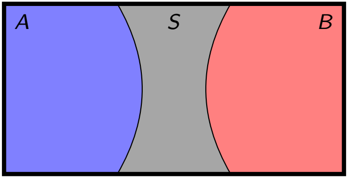


$$\mathcal{C} = \mathcal{C}_A \sqcup \mathcal{C}_B \sqcup \mathcal{C}_S$$

$$\text{(HC)} \quad p(\boldsymbol{x}) = \textstyle\prod_{C \in \mathcal{C}} V_C(\boldsymbol{x}_C)$$

$$p(\underbrace{\boldsymbol{x}_A, \boldsymbol{x}_B, \boldsymbol{x}_S}_{\boldsymbol{x}}) = \prod_{C \in \mathcal{C}_A} V_C(\boldsymbol{x}_C) \cdot \prod_{C \in \mathcal{C}_B} V_C(\boldsymbol{x}_C) \cdot \prod_{C \in \mathcal{C}_S} V_C(\boldsymbol{x}_C)$$

Przypadek I. $A \cup B \cup S = V$

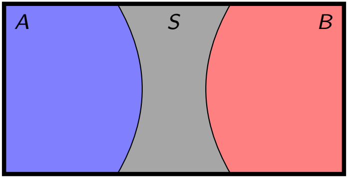


$$\mathcal{C} = \mathcal{C}_A \sqcup \mathcal{C}_B \sqcup \mathcal{C}_S$$

$$\text{(HC)} \quad p(\boldsymbol{x}) = \prod_{C \in \mathcal{C}} V_C(\boldsymbol{x}_C)$$

$$p(\underbrace{\boldsymbol{x}_A, \boldsymbol{x}_B, \boldsymbol{x}_S}_{\boldsymbol{x}}) = \underbrace{\prod_{C \in \mathcal{C}_A} V_C(\boldsymbol{x}_C)}_{\Phi(\boldsymbol{x}_A, \boldsymbol{x}_S)} \cdot \prod_{C \in \mathcal{C}_B} V_C(\boldsymbol{x}_C) \cdot \prod_{C \in \mathcal{C}_S} V_C(\boldsymbol{x}_C)$$

Przypadek I. $A \cup B \cup S = V$



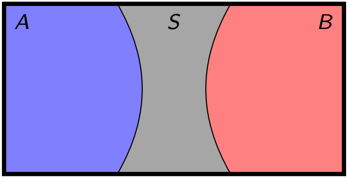
$$\mathcal{C} = \mathcal{C}_A \sqcup \mathcal{C}_B \sqcup \mathcal{C}_S$$

$$\text{(HC)} \quad p(\boldsymbol{x}) = \textstyle\prod_{C \in \mathcal{C}} V_C(\boldsymbol{x}_C)$$

$$p(\underbrace{\boldsymbol{x}_A, \boldsymbol{x}_B, \boldsymbol{x}_S}_{\boldsymbol{x}}) = \underbrace{\prod_{C \in \mathcal{C}_A} V_C(\boldsymbol{x}_C)}_{\Phi(\boldsymbol{x}_A, \boldsymbol{x}_S)} \cdot \underbrace{\prod_{C \in \mathcal{C}_B} V_C(\boldsymbol{x}_C) \cdot \prod_{C \in \mathcal{C}_S} V_C(\boldsymbol{x}_C)}_{\Psi(\boldsymbol{x}_B, \boldsymbol{x}_S)}$$

Dowód $\Leftarrow$

Przypadek I. $A \cup B \cup S = V$



$$\mathcal{C} = \mathcal{C}_A \sqcup \mathcal{C}_B \sqcup \mathcal{C}_S$$

$$\text{(HC)} \quad p(\boldsymbol{x}) = \prod_{C \in \mathcal{C}} V_C(\boldsymbol{x}_C)$$

$$p(\underbrace{\boldsymbol{x}_A, \boldsymbol{x}_B, \boldsymbol{x}_S}_{\boldsymbol{x}}) = \underbrace{\prod_{C \in \mathcal{C}_A} V_C(\boldsymbol{x}_C)}_{\Phi(\boldsymbol{x}_A, \boldsymbol{x}_S)} \cdot \underbrace{\prod_{C \in \mathcal{C}_B} V_C(\boldsymbol{x}_C) \cdot \prod_{C \in \mathcal{C}_S} V_C(\boldsymbol{x}_C)}_{\Psi(\boldsymbol{x}_B, \boldsymbol{x}_S)}$$

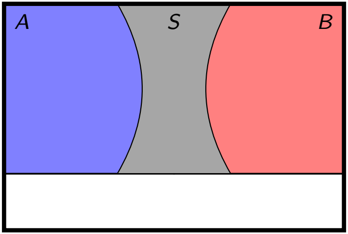
… a zatem $\boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid \boldsymbol{X}_S$

Dowód $\Leftarrow$

Przypadek II. $A \cup B \cup S \neq V$

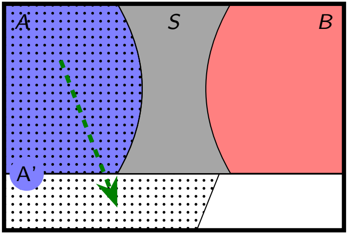
Dowód $\Leftarrow$

Przypadek II. $A \cup B \cup S \neq V$



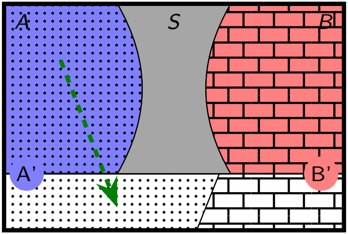
Dowód $\Leftarrow$

Przypadek II. $A \cup B \cup S \neq V$



A': połączone z A ścieżką poza S

Przypadek II. $A \cup B \cup S \neq V$

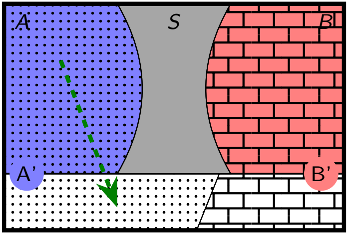


A': połączone z A ścieżką poza S

$B' = V \setminus (A' \cup S)$

Dowód $\Leftarrow$

Przypadek II. $A \cup B \cup S \neq V$



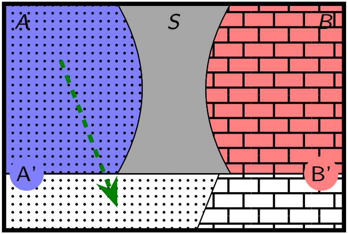
A': połączone z A ścieżką poza S

$B' = V \setminus (A' \cup S)$

Wówczas nie można przejść z A' do B' omijając S.

Dowód $\Leftarrow$

Przypadek II. $A \cup B \cup S \neq V$



A': połączone z A ścieżką poza S

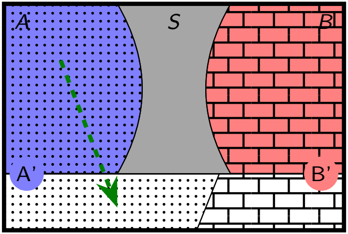
$B' = V \setminus (A' \cup S)$

Wówczas nie można przejść z A' do B' omijając S.

Ponadto $V = A' \cup B' \cup S$

Dowód $\Leftarrow$

Przypadek II. $A \cup B \cup S \neq V$



A': połączone z A ścieżką poza S

$B' = V \setminus (A' \cup S)$

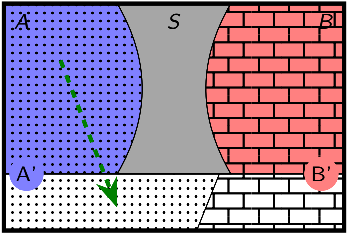
Wówczas nie można przejść z A' do B' omijając S.

Ponadto $V = A' \cup B' \cup S$

Z wcześniejszych rozważań wynika $\boldsymbol{X}_{A'} \perp\!\!\!\perp \boldsymbol{X}_{B'} \mid \boldsymbol{X}_S$

Dowód $\Leftarrow$

Przypadek II. $A \cup B \cup S \neq V$



A': połączone z A ścieżką poza S

$B' = V \setminus (A' \cup S)$

Wówczas nie można przejść z A' do B' omijając S.

Ponadto $V = A' \cup B' \cup S$

Z wcześniejszych rozważań wynika $\boldsymbol{X}_{A'} \perp\!\!\!\perp \boldsymbol{X}_{B'} \mid \boldsymbol{X}_S$

Ponieważ $A \subseteq A'$ i $B \subseteq B'$, więc $\boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid \boldsymbol{X}_S$ $\square$

Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$.

Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$.

Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$. Ponieważ W jest samorodny

Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$. Ponieważ W jest samorodny więc

$$p(\boldsymbol{x}_W) = \prod_{i \in W} p(\, x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)} \,)$$

Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$. Ponieważ W jest samorodny więc

$$p(\boldsymbol{x}_W) = \prod_{i \in W} p(\underbrace{x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)}}_{\text{klika w } \mathcal{G}_W^{\mathrm{m}}})$$

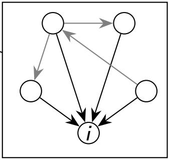
Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$. Ponieważ W jest samorodny więc

$$p(\boldsymbol{x}_W) = \prod_{i \in W} p(\underbrace{x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)}}_{\text{klika w } \mathcal{G}_W^{\mathrm{m}}})$$



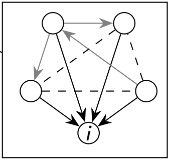
Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$. Ponieważ W jest samorodny więc

$$p(\boldsymbol{x}_W) = \prod_{i \in W} p(\underbrace{x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)}}_{\text{klika w } \mathcal{G}_W^{\mathrm{m}}})$$



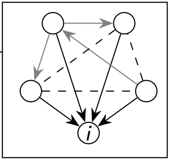
Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$. Ponieważ W jest samorodny więc

$$p(\boldsymbol{x}_W) = \prod_{i \in W} p(\underbrace{x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)}}_{\text{klika w } \mathcal{G}_W^{\mathrm{m}}})$$



Zmienne $\boldsymbol{X}_W$ spełniają zatem warunek (HC) względem grafu $\mathcal{G}_W^{\mathrm{m}}$.

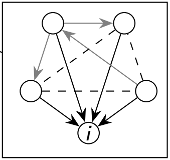
Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$. Ponieważ W jest samorodny więc

$$p(\boldsymbol{x}_W) = \prod_{i \in W} p(\underbrace{x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)}}_{\text{klika w } \mathcal{G}^{\mathrm{m}}_W})$$



Zmienne $\boldsymbol{X}_W$ spełniają zatem warunek (HC) względem grafu $\mathcal{G}^{\mathrm{m}}_W$.

Czyli $A \not\!\!-\!\!- B \mid S$ w $\mathcal{G}^{\mathrm{m}}_W$ implikuje $\boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid \boldsymbol{X}_S$.

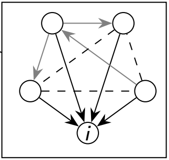
Dowód twierdzenia Pearla

Niech $\mathcal{G}$ będzie siecią bayesowską dla zmiennych $X_1, \ldots, X_n$. Mamy

$$p(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)})$$

Niech $W = \mathrm{An}(A \cup B \cup S)$. Ponieważ W jest samorodny więc

$$p(\boldsymbol{x}_W) = \prod_{i \in W} p(\underbrace{x_i \mid \boldsymbol{x}_{\mathrm{pa}(i)}}_{\text{klika w } \mathcal{G}^{\mathrm{m}}_W})$$



Zmienne $\boldsymbol{X}_W$ spełniają zatem warunek (HC) względem grafu $\mathcal{G}^{\mathrm{m}}_W$.

Czyli $A \not\!\!- B \mid S$ w $\mathcal{G}^{\mathrm{m}}_W$ implikuje $\boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid \boldsymbol{X}_S$.

Zatem $A \not\to B \mid S$ w $\mathcal{G}$ implikuje $\boldsymbol{X}_A \perp\!\!\!\perp \boldsymbol{X}_B \mid \boldsymbol{X}_S$. $\square$

Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*

Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)

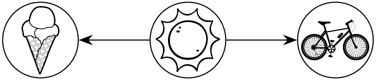
Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)
  - teoria sieci bayesowskich

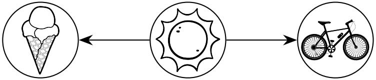
Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)
  - teoria sieci bayesowskich
  - **teoria przyczynowości**

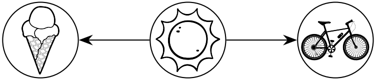
Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)
  - teoria sieci bayesowskich
  - **teoria przyczynowości**
- 

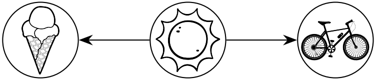
Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)
  - teoria sieci bayesowskich
  - **teoria przyczynowości**
- 
  - kupno loda nie zwiększa szans na ładną pogodę!

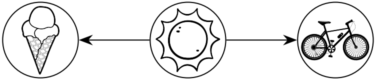
Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)
  - teoria sieci bayesowskich
  - **teoria przyczynowości**
- 
  - kupno loda nie zwiększa szans na ładną pogodę!
  - warunkowanie przez interwencję $\neq$ warunkowanie przez obserwację

Kilka komentarzy na koniec

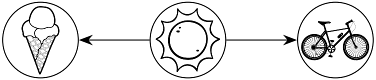
- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)
  - teoria sieci bayesowskich
  - **teoria przyczynowości**
- 
  - kupno loda nie zwiększa szans na ładną pogodę!
  - warunkowanie przez interwencję $\neq$ warunkowanie przez obserwację
  - kiedy można pr. interwencyjne obliczyć na podstawie obserwacyjnych?

Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)
  - teoria sieci bayesowskich
  - **teoria przyczynowości**
- 
  - kupno loda nie zwiększa szans na ładną pogodę!
  - warunkowanie przez interwencję $\neq$ warunkowanie przez obserwację
  - kiedy można pr. interwencyjne obliczyć na podstawie obserwacyjnych?
- zainteresowanym polecam J. Pearl „Przyczyny i skutki"

Kilka komentarzy na koniec

- *Każdy z nas, matematyków, ma w pamięci pytanie, które zapoczątkowało uprawianą przez nas problematykę (...)*
- rozważane przez nas pytanie leży u podstaw dwóch dziedzin (zapoczątkowanych przez J. Pearla)
  - teoria sieci bayesowskich
  - **teoria przyczynowości**

- 
  - kupno loda nie zwiększa szans na ładną pogodę!
  - warunkowanie przez interwencję $\neq$ warunkowanie przez obserwację
  - kiedy można pr. interwencyjne obliczyć na podstawie obserwacyjnych?
- zainteresowanym polecam J. Pearl „Przyczyny i skutki”

Dziękuję za uwagę!