

Úvod do bayesovských sítí

Jiří Vomlel

Ústav teorie informace a automatizace
Akademie věd České republiky
<http://www.utia.cz/vomlel>

30. října 2008

- Velkými písmeny budeme označovat veličiny, např. *Horečka*.

Podmíněná pravděpodobnost

- Velkými písmeny budeme označovat veličiny, např. *Horečka*.
- Malými písmeny stavy veličin, např. *ano*, *ne*.

Podmíněná pravděpodobnost

- Velkými písmeny budeme označovat veličiny, např. *Horečka*.
- Malými písmeny stavy veličin, např. *ano*, *ne*.
- Budeme se zabývat pravděpodobnostmi různých jevů, např. jevu *Horečka=ano*, kterou označíme $P(\text{Horečka}=\text{ano})$, nebo jevu *Horečka=ano a Angína=ano*, označíme $P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano})$

Podmíněná pravděpodobnost

- Velkými písmeny budeme označovat veličiny, např. *Horečka*.
- Malými písmeny stavy veličin, např. *ano*, *ne*.
- Budeme se zabývat pravděpodobnostmi různých jevů, např. jevu *Horečka=ano*, kterou označíme $P(\text{Horečka}=\text{ano})$, nebo jevu *Horečka=ano a Angína=ano*, označíme $P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano})$
- Podmíněná pravděpodobnost jevu *Horečka=ano* při pozorovaném jevu *Angína=ano*, kterou označíme $P(\text{Horečka}=\text{ano} \mid \text{Angína}=\text{ano})$, je pravděpodobnost splňující vztah

$$\begin{aligned} P(\text{Horečka}=\text{ano} \mid \text{Angína}=\text{ano}) \cdot P(\text{Angína}=\text{ano}) \\ = P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano}) . \end{aligned}$$

Podmíněná pravděpodobnost

- Velkými písmeny budeme označovat veličiny, např. *Horečka*.
- Malými písmeny stavy veličin, např. *ano*, *ne*.
- Budeme se zabývat pravděpodobnostmi různých jevů, např. jevu *Horečka=ano*, kterou označíme $P(\text{Horečka}=\text{ano})$, nebo jevu *Horečka=ano a Angína=ano*, označíme $P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano})$
- Podmíněná pravděpodobnost jevu *Horečka=ano* při pozorovaném jevu *Angína=ano*, kterou označíme $P(\text{Horečka}=\text{ano} \mid \text{Angína}=\text{ano})$, je pravděpodobnost splňující vztah

$$\begin{aligned} P(\text{Horečka}=\text{ano} \mid \text{Angína}=\text{ano}) \cdot P(\text{Angína}=\text{ano}) \\ = P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano}) . \end{aligned}$$

- Jestliže $P(\text{Angína}=\text{ano})$ je nenulová, pak

$$P(\text{Horečka}=\text{ano} \mid \text{Angína}=\text{ano}) = \frac{P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano})}{P(\text{Angína}=\text{ano})} .$$

- Pozorováním jsme získali následující pravděpodobnosti jevů:

$$P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano}) = 0.015$$

$$P(\text{Horečka}=\text{ne}, \text{Angína}=\text{ano}) = 0.005$$

$$P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ne}) = 0.08$$

$$P(\text{Horečka}=\text{ne}, \text{Angína}=\text{ne}) = 0.90$$

Podmíněná pravděpodobnost - příklad

- Pozorováním jsme získali následující pravděpodobnosti jevů:

$$P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano}) = 0.015$$

$$P(\text{Horečka}=\text{ne}, \text{Angína}=\text{ano}) = 0.005$$

$$P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ne}) = 0.08$$

$$P(\text{Horečka}=\text{ne}, \text{Angína}=\text{ne}) = 0.90$$

- Můžeme spočítat např.

$$P(\text{Horečka}=\text{ano} \mid \text{Angína}=\text{ano}) = \frac{0.015}{0.015 + 0.005} = \frac{0.015}{0.020} = \frac{3}{4}$$

Podmíněná pravděpodobnost - příklad

- Pozorováním jsme získali následující pravděpodobnosti jevů:

$$P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ano}) = 0.015$$

$$P(\text{Horečka}=\text{ne}, \text{Angína}=\text{ano}) = 0.005$$

$$P(\text{Horečka}=\text{ano}, \text{Angína}=\text{ne}) = 0.08$$

$$P(\text{Horečka}=\text{ne}, \text{Angína}=\text{ne}) = 0.90$$

- Můžeme spočítat např.

$$P(\text{Horečka}=\text{ano} \mid \text{Angína}=\text{ano}) = \frac{0.015}{0.015 + 0.005} = \frac{0.015}{0.020} = \frac{3}{4}$$

- nebo

$$P(\text{Angína}=\text{ano} \mid \text{Horečka}=\text{ano}) = \frac{0.015}{0.015 + 0.08} = \frac{0.015}{0.095} = \frac{3}{19}$$

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

$$\begin{aligned}P(A = a|B = b) &= \frac{P(A = a, B = b)}{P(B = b)} \\ &= \frac{P(B = b|A = a) \cdot P(A = a)}{\sum_i P(A = i, B = b)}\end{aligned}$$

$$\begin{aligned}P(A = a|B = b) &= \frac{P(A = a, B = b)}{P(B = b)} \\&= \frac{P(B = b|A = a) \cdot P(A = a)}{\sum_i P(A = i, B = b)} \\&= \frac{P(B = b|A = a) \cdot P(A = a)}{\sum_i P(B = b|A = i) \cdot P(A = i)}\end{aligned}$$

Bayesův vzorec - příklad

- D ... v noci přšelo, s hodnotami $a = ano$ a $n = ne$.

Bayesův vzorec - příklad

- D ... v noci pršelo, s hodnotami $a = ano$ a $n = ne$.
- M ... tráva je mokrá, s hodnotami $a = ano$ a $n = ne$.

Bayesův vzorec - příklad

- D ... v noci pršelo, s hodnotami $a = ano$ a $n = ne$.
- M ... tráva je mokrá, s hodnotami $a = ano$ a $n = ne$.
- Jestliže v noci pršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = a) = \frac{4}{5}$.

Bayesův vzorec - příklad

- D ... v noci pršelo, s hodnotami $a = ano$ a $n = ne$.
- M ... tráva je mokrá, s hodnotami $a = ano$ a $n = ne$.
- Jestliže v noci pršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = a) = \frac{4}{5}$.
- Jestliže v noci nepršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = n) = \frac{2}{5}$.

Bayesův vzorec - příklad

- D ... v noci pršelo, s hodnotami $a = ano$ a $n = ne$.
- M ... tráva je mokrá, s hodnotami $a = ano$ a $n = ne$.
- Jestliže v noci pršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = a) = \frac{4}{5}$.
- Jestliže v noci nepršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = n) = \frac{2}{5}$.
- Pravděpodobnost, že v noci prší, je $P(D = a) = \frac{1}{4}$.

Bayesův vzorec - příklad

- D ... v noci pršelo, s hodnotami $a = ano$ a $n = ne$.
- M ... tráva je mokrá, s hodnotami $a = ano$ a $n = ne$.
- Jestliže v noci pršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = a) = \frac{4}{5}$.
- Jestliže v noci nepršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = n) = \frac{2}{5}$.
- Pravděpodobnost, že v noci prší, je $P(D = a) = \frac{1}{4}$.
- Ráno vidíme, že tráva je mokrá. Jaká je pravděpodobnost, že v noci pršelo?

Bayesův vzorec - příklad

- D ... v noci pršelo, s hodnotami $a = ano$ a $n = ne$.
- M ... tráva je mokrá, s hodnotami $a = ano$ a $n = ne$.
- Jestliže v noci pršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = a) = \frac{4}{5}$.
- Jestliže v noci nepršelo, pak je tráva mokrá s pravděpodobností $P(M = a|D = n) = \frac{2}{5}$.
- Pravděpodobnost, že v noci prší, je $P(D = a) = \frac{1}{4}$.
- Ráno vidíme, že tráva je mokrá. Jaká je pravděpodobnost, že v noci pršelo?

$$P(D = a|M = a)$$

$$= \frac{P(M = a|D = a) \cdot P(D = a)}{P(M = a|D = a) \cdot P(D = a) + P(M = a|D = n) \cdot P(D = n)}$$

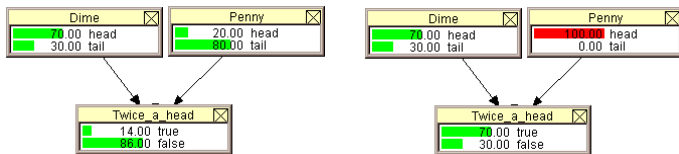
$$= \frac{\frac{4}{5} \cdot \frac{1}{4}}{\frac{4}{5} \cdot \frac{1}{4} + \frac{2}{5} \cdot \frac{3}{4}} = \frac{\frac{1}{5}}{\frac{1}{5} + \frac{6}{20}} = \frac{\frac{1}{5}}{\frac{1}{2}} = \frac{2}{5}$$

Nezávislost dvou veličin

Mějme dvě velmi podivné mince *Dime* and *Penny*. *Dime* má pravděpodobnost, že padne *head* 0.7, ale *Penny* pouze 0.2.

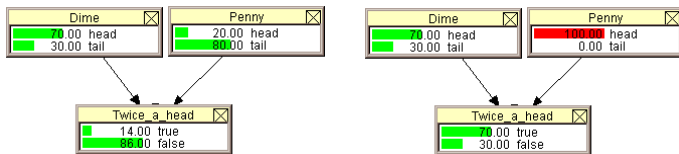
Nezávislost dvou veličin

Mějme dvě velmi podivné mince *Dime* and *Penny*. *Dime* má pravděpodobnost, že padne *head* 0.7, ale *Penny* pouze 0.2.



Nezávislost dvou veličin

Mějme dvě velmi podivné mince *Dime* and *Penny*. *Dime* má pravděpodobnost, že padne *head* 0.7, ale *Penny* pouze 0.2.

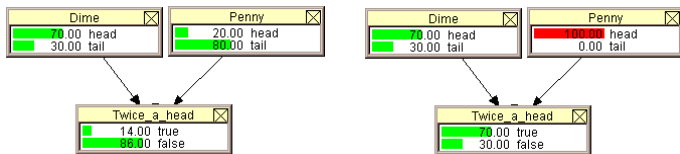


Pravděpodobnost společného výskytu hodnot veličin je rovna součinu pravděpodobností jednotlivých pravděpodobností.

$$\begin{aligned} P(Dime = head, Penny = head) \\ = P(Dime = head) \cdot P(Penny = head) \end{aligned}$$

Nezávislost dvou veličin

Mějme dvě velmi podivné mince *Dime* and *Penny*. *Dime* má pravděpodobnost, že padne *head* 0.7, ale *Penny* pouze 0.2.



Pravděpodobnost společného výskytu hodnot veličin je rovna součinu pravděpodobností jednotlivých pravděpodobností.

$$\begin{aligned} P(Dime = head, Penny = head) \\ = P(Dime = head) \cdot P(Penny = head) \end{aligned}$$

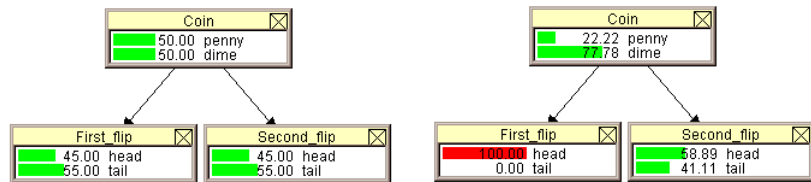
Též, zjistíme-li hodnotu jedné veličiny, nemá to vliv na hodnotu druhé veličiny.

$$P(Dime = head | Penny = head) = P(Dime = head)$$

Kolega náhodně vybere jednu z těch dvou podivných mincí a udělá s ní nejprve jeden a pak druhý hod. My nevíme, jakou minci si vybral.

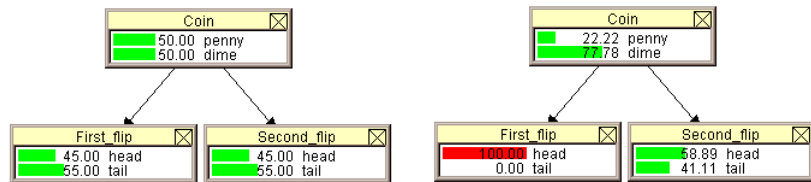
Závislost dvou veličin

Kolega náhodně vybere jednu z těchto dvou podivných mincí a udělá s ní nejprve jeden a pak druhý hod. My nevíme, jakou minci si vybral.



Závislost dvou veličin

Kolega náhodně vybere jednu z těch dvou podivných mincí a udělá s ní nejprve jeden a pak druhý hod. My nevíme, jakou minci si vybral.

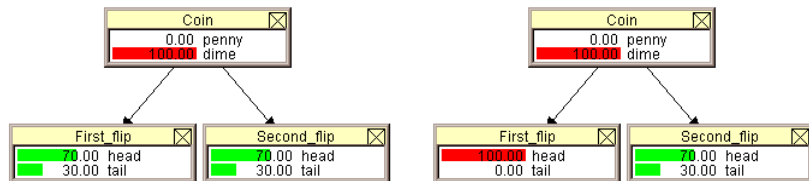


Výsledek **prvního** hodu **má vliv** na pravděpodobnost výsledku **druhého** hodu.

Nyní předpokládejme, že víme, kterou minci si kolega vybral.

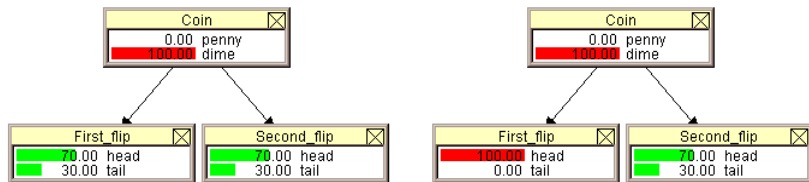
Podmíněná nezávislost dvou veličin

Nyní předpokládejme, že víme, kterou minci si kolega vybral.



Podmíněná nezávislost dvou veličin

Nyní předpokládejme, že víme, kterou minci si kolega vybral.



Jestliže víme, **která mince** byla použita, pak výsledek **prvního** hodů **nemá vliv** na pravděpodobnost výsledku **druhého** hodů.

Pravděpodobnost společného výskytu hodnot veličin při dané hodnotě třetí veličiny je rovna součinu pravděpodobností jednotlivých podmíněných pravděpodobností:

$$\begin{aligned} &P(\textit{First_flip} = \textit{head}, \textit{Second_flip} = \textit{head} | \textit{Coin} = \textit{dime}) \\ &= P(\textit{First_flip} = \textit{head} | \textit{Coin} = \textit{dime}) \\ &\quad \cdot P(\textit{Second_flip} = \textit{head} | \textit{Coin} = \textit{dime}) \end{aligned}$$

Podmíněná nezávislost

Pravděpodobnost společného výskytu hodnot veličin při dané hodnotě třetí veličiny je rovna součinu pravděpodobností jednotlivých podmíněných pravděpodobností:

$$\begin{aligned} &P(\textit{First_flip} = \textit{head}, \textit{Second_flip} = \textit{head} | \textit{Coin} = \textit{dime}) \\ &= P(\textit{First_flip} = \textit{head} | \textit{Coin} = \textit{dime}) \\ &\quad \cdot P(\textit{Second_flip} = \textit{head} | \textit{Coin} = \textit{dime}) \end{aligned}$$

Jestliže známe minci, pak výsledek **prvního** hodů **nemá vliv** na pravděpodobnost výsledku **druhého** hodů.

$$\begin{aligned} &P(\textit{Second_flip} = \textit{head} | \textit{Coin} = \textit{dime}, \textit{First_flip} = \textit{head}) \\ &= P(\textit{Second_flip} = \textit{head} | \textit{Coin} = \textit{dime}) \end{aligned}$$

Z definice podmíněné pravděpodobnosti plyne, že můžeme psát:

$$P(A, B, C, D) = P(A|B, C, D) \cdot P(B, C, D)$$

Z definice podmíněné pravděpodobnosti plyne, že můžeme psát:

$$\begin{aligned}P(A, B, C, D) &= P(A|B, C, D) \cdot P(B, C, D) \\ &= P(A|B, C, D) \cdot P(B|C, D) \cdot P(C, D)\end{aligned}$$

Z definice podmíněné pravděpodobnosti plyne, že můžeme psát:

$$\begin{aligned}P(A, B, C, D) &= P(A|B, C, D) \cdot P(B, C, D) \\ &= P(A|B, C, D) \cdot P(B|C, D) \cdot P(C, D) \\ &= P(A|B, C, D) \cdot P(B|C, D) \cdot P(C|D) \cdot P(D)\end{aligned}$$

Proč je Holmesův trávník mokrý?

- Máme 4 veličiny:

<i>Holm</i>	Je Holmesův trávník mokrý?
<i>Rn</i>	Pršelo v noci?
<i>Sprnk</i>	Byl Holmesův postřikovač zapnutý?
<i>Wat</i>	Je Watsonův trávník mokrý?

Proč je Holmesův trávník mokrý?

- Máme 4 veličiny:

<i>Holm</i>	Je Holmesův trávník mokrý?
<i>Rn</i>	Pršelo v noci?
<i>Sprnk</i>	Byl Holmesův postřikovač zapnutý?
<i>Wat</i>	Je Watsonův trávník mokrý?
- Holmesův trávník může být mokrý, buď protože pršelo, nebo protože měl zapnutý postřikovač.

Proč je Holmesův trávník mokrý?

- Máme 4 veličiny:

<i>Holm</i>	Je Holmesův trávník mokrý?
<i>Rn</i>	Pršelo v noci?
<i>Sprnk</i>	Byl Holmesův postřikovač zapnutý?
<i>Wat</i>	Je Watsonův trávník mokrý?
- Holmesův trávník může být mokrý, buď protože pršelo, nebo protože měl zapnutý postřikovač.
- Watsonův trávník může být mokrý, protože pršelo. Holmesův postřikovač nemá vliv na Watsonův trávník.

Proč je Holmesův trávník mokrý?

- Máme 4 veličiny:

<i>Holm</i>	Je Holmesův trávník mokrý?
<i>Rn</i>	Pršelo v noci?
<i>Sprnk</i>	Byl Holmesův postřikovač zapnutý?
<i>Wat</i>	Je Watsonův trávník mokrý?
- Holmesův trávník může být mokrý, buď protože pršelo, nebo protože měl zapnutý postřikovač.
- Watsonův trávník může být mokrý, protože pršelo. Holmesův postřikovač nemá vliv na Watsonův trávník.
- Déšť nesouvisí s tím, jestli má Holmes zapnutý postřikovač.

Proč je Holmesův trávník mokrý?

- Máme 4 veličiny:

<i>Holm</i>	Je Holmesův trávník mokrý?
<i>Rn</i>	Pršelo v noci?
<i>Sprnk</i>	Byl Holmesův postřikovač zapnutý?
<i>Wat</i>	Je Watsonův trávník mokrý?
- Holmesův trávník může být mokrý, buď protože pršelo, nebo protože měl zapnutý postřikovač.
- Watsonův trávník může být mokrý, protože pršelo. Holmesův postřikovač nemá vliv na Watsonův trávník.
- Déšť nesouvisí s tím, jestli má Holmes zapnutý postřikovač.
- **Řetězcové pravidlo** a **podmíněné nezávislosti** dávají:

Proč je Holmesův trávník mokrý?

- Máme 4 veličiny:

<i>Holm</i>	Je Holmesův trávník mokrý?
<i>Rn</i>	Pršelo v noci?
<i>Sprnk</i>	Byl Holmesův postřikovač zapnutý?
<i>Wat</i>	Je Watsonův trávník mokrý?
- Holmesův trávník může být mokrý, buď protože pršelo, nebo protože měl zapnutý postřikovač.
- Watsonův trávník může být mokrý, protože pršelo. Holmesův postřikovač nemá vliv na Watsonův trávník.
- Déšť nesouvisí s tím, jestli má Holmes zapnutý postřikovač.
- **Řetězcové pravidlo** a **podmínné nezávislosti** dávají:

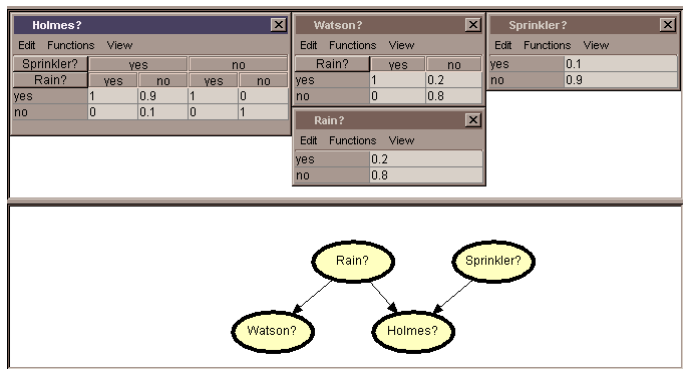
$$\begin{aligned} &P(\textit{Holm}, \textit{Wat}, \textit{Rn}, \textit{Sprnk}) \\ &= P(\textit{Holm} | \textit{Wat}, \textit{Rn}, \textit{Sprnk}) \cdot P(\textit{Wat} | \textit{Rn}, \textit{Sprnk}) \\ &\quad \cdot P(\textit{Rn} | \textit{Sprnk}) \cdot P(\textit{Sprnk}) \\ &= P(\textit{Holm} | \textit{Rn}, \textit{Sprnk}) \cdot P(\textit{Wat} | \textit{Rn}) \cdot P(\textit{Rn}) \cdot P(\textit{Sprnk}) \end{aligned}$$

Proč je Holmesův trávník mokrý?

$$\begin{aligned} &P(\text{Holm}, \text{Wat}, \text{Rn}, \text{Sprnk}) \\ &= P(\text{Holm}|\text{Rn}, \text{Sprnk}) \cdot P(\text{Wat}|\text{Rn}) \cdot P(\text{Rn}) \cdot P(\text{Sprnk}) \end{aligned}$$

Proč je Holmesův trávník mokrý?

$$P(\text{Holm}, \text{Wat}, \text{Rn}, \text{Sprnk}) \\ = P(\text{Holm} | \text{Rn}, \text{Sprnk}) \cdot P(\text{Wat} | \text{Rn}) \cdot P(\text{Rn}) \cdot P(\text{Sprnk})$$



- acyklický orientovaný graf (DAG) $G = (V, E)$

- acyklický orientovaný graf (DAG) $G = (V, E)$
- každý uzel $i \in V$ odpovídá jedné náhodné veličině X_i s konečným počtem navzájem disjunktních hodnot \mathbb{X}_i

- acyklický orientovaný graf (DAG) $G = (V, E)$
- každý uzel $i \in V$ odpovídá jedné náhodné veličině X_i s konečným počtem navzájem disjunktních hodnot \mathbb{X}_i
- acyklický orientovaný graf (DAG) reprezentuje podmíněně nezávislostní vztahy mezi veličinami $(X_i)_{i \in V}$

Definice bayesovské sítě

- acyklický orientovaný graf (DAG) $G = (V, E)$
- každý uzel $i \in V$ odpovídá jedné náhodné veličině X_i s konečným počtem navzájem disjunktních hodnot \mathbb{X}_i
- acyklický orientovaný graf (DAG) reprezentuje podmíněně nezávislostní vztahy mezi veličinami $(X_i)_{i \in V}$
- $pa(i)$ bude označovat množinu rodičů uzlu i v grafu G

Definice bayesovské sítě

- acyklický orientovaný graf (DAG) $G = (V, E)$
- každý uzel $i \in V$ odpovídá jedné náhodné veličině X_i s konečným počtem navzájem disjunktních hodnot \mathbb{X}_i
- acyklický orientovaný graf (DAG) reprezentuje podmíněně nezávislostní vztahy mezi veličinami $(X_i)_{i \in V}$
- $pa(i)$ bude označovat množinu rodičů uzlu i v grafu G
- ke každému uzlu $i \in V$ odpovídá podmíněná pravděpodobnostní distribuce $P(X_i \mid (X_j)_{j \in pa(i)})$

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo
- cesta obsahuje uzel, ve kterém se hrany **setkávají** “head-to-head” a ani on, ani žádný jeho následník **nenáleží** do \mathcal{Y} .

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo
- cesta obsahuje uzel, ve kterém se hrany **setkávají** “head-to-head” a ani on, ani žádný jeho následník **nenáleží** do \mathcal{Y} .

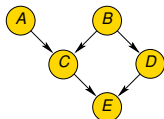
Jestliže i a j jsou d-separovány množinou uzlů \mathcal{Y} , pak veličiny X_i a X_j jsou nezávislé dáno $(X_k)_{k \in \mathcal{Y}}$.

Podmíněné nezávislost dané grafem - d-separace

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo
- cesta obsahuje uzel, ve kterém se hrany **setkávají** “head-to-head” a ani on, ani žádný jeho následník **nenáleží** do \mathcal{Y} .

Jestliže i a j jsou d-separovány množinou uzlů \mathcal{Y} , pak veličiny X_i a X_j jsou nezávislé dáno $(X_k)_{k \in \mathcal{Y}}$.

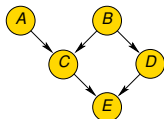


Podmíněné nezávislost dané grafem - d-separace

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo
- cesta obsahuje uzel, ve kterém se hrany **setkávají** “head-to-head” a ani on, ani žádný jeho následník **nenáleží** do \mathcal{Y} .

Jestliže i a j jsou d-separovány množinou uzlů \mathcal{Y} , pak veličiny X_i a X_j jsou nezávislé dáno $(X_k)_{k \in \mathcal{Y}}$.



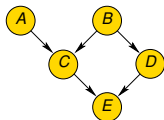
$$A \perp\!\!\!\perp D$$

Podmíněné nezávislost dané grafem - d-separace

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo
- cesta obsahuje uzel, ve kterém se hrany **setkávají** “head-to-head” a ani on, ani žádný jeho následník **nenáleží** do \mathcal{Y} .

Jestliže i a j jsou d-separovány množinou uzlů \mathcal{Y} , pak veličiny X_i a X_j jsou nezávislé dáno $(X_k)_{k \in \mathcal{Y}}$.



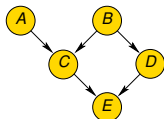
$$A \perp\!\!\!\perp D \mid B$$

Podmíněné nezávislost dané grafem - d-separace

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo
- cesta obsahuje uzel, ve kterém se hrany **setkávají** “head-to-head” a ani on, ani žádný jeho následník **nenáleží** do \mathcal{Y} .

Jestliže i a j jsou d-separovány množinou uzlů \mathcal{Y} , pak veličiny X_i a X_j jsou nezávislé dáno $(X_k)_{k \in \mathcal{Y}}$.



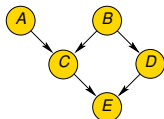
$$A \not\perp D \mid B, E$$

Podmíněné nezávislost dané grafem - d-separace

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo
- cesta obsahuje uzel, ve kterém se hrany **setkávají** “head-to-head” a ani on, ani žádný jeho následník **nenáleží** do \mathcal{Y} .

Jestliže i a j jsou d-separovány množinou uzlů \mathcal{Y} , pak veličiny X_i a X_j jsou nezávislé dáno $(X_k)_{k \in \mathcal{Y}}$.



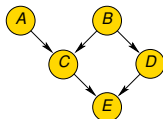
$C?D \mid E$

Podmíněné nezávislost dané grafem - d-separace

Dva uzly i a j jsou d-separovány množinou uzlů \mathcal{Y} , jestliže pro všechny cesty mezi i a j platí:

- cesta obsahuje uzel, ve kterém se hrany **nesetkávají** “head-to-head” a který **náleží** do \mathcal{Y} nebo
- cesta obsahuje uzel, ve kterém se hrany **setkávají** “head-to-head” a ani on, ani žádný jeho následník **nenáleží** do \mathcal{Y} .

Jestliže i a j jsou d-separovány množinou uzlů \mathcal{Y} , pak veličiny X_i a X_j jsou nezávislé dáno $(X_k)_{k \in \mathcal{Y}}$.



$$C \not\perp D \mid E$$

Použijeme-li **řetězcové pravidlo** dostaneme:

$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i \mid X_{i-1}, \dots, X_1)$$

Použijeme-li **řetězcové pravidlo** dostaneme:

$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i \mid X_{i-1}, \dots, X_1)$$

Použijeme-li podmíněné nezávislosti z grafu pak dostaneme

$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i \mid (X_j)_{j \in pa(i)})$$

Pravděpodobnostní distribuce bayesovské sítě

Použijeme-li **řetězcové pravidlo** dostaneme:

$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i \mid X_{i-1}, \dots, X_1)$$

Použijeme-li podmíněné nezávislosti z grafu pak dostaneme

$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i \mid (X_j)_{j \in pa(i)})$$

Toto je pravděpodobnostní distribuce representovaná **bayesovskou sítí**.

Pravděpodobnostní distribuce bayesovské sítě

Použijeme-li **řetězcové pravidlo** dostaneme:

$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i | X_{i-1}, \dots, X_1)$$

Použijeme-li podmíněné nezávislosti z grafu pak dostaneme

$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i | (X_j)_{j \in pa(i)})$$

Toto je pravděpodobnostní distribuce representovaná **bayesovskou sítí**.

Pro tuto distribuci platí:

- splňuje podmíněné nezávislosti zakódované v grafu a

Pravděpodobnostní distribuce bayesovské sítě

Použijeme-li **řetězcové pravidlo** dostaneme:

$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i \mid X_{i-1}, \dots, X_1)$$

Použijeme-li podmíněné nezávislosti z grafu pak dostaneme

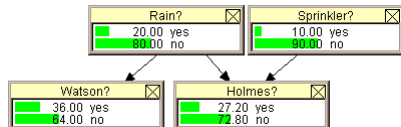
$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i \mid (X_j)_{j \in pa(i)})$$

Toto je pravděpodobnostní distribuce reprezentovaná **bayesovskou sítí**.

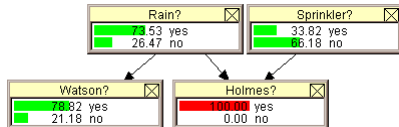
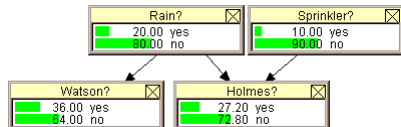
Pro tuto distribuci platí:

- splňuje podmíněné nezávislosti zakódované v grafu a
- její podmíněné pravděpodobnostní distribuce odpovídají $P(X_i \mid (X_j)_{j \in pa(i)})$, $i \in V$.

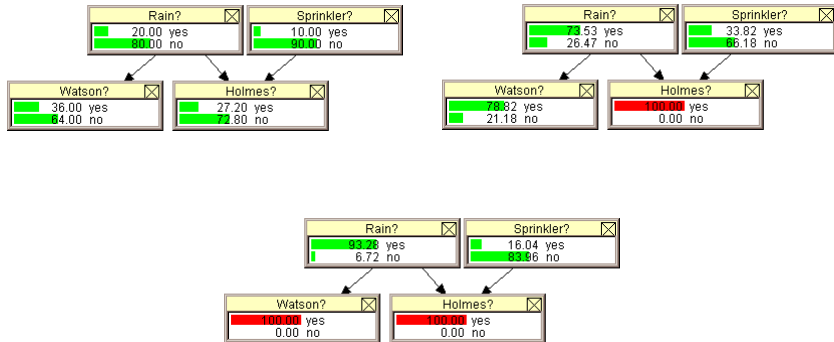
Byl Holmesův postřikovač zapnutý?



Byl Holmesův postřikovač zapnutý?



Byl Holmesův postřikovač zapnutý?



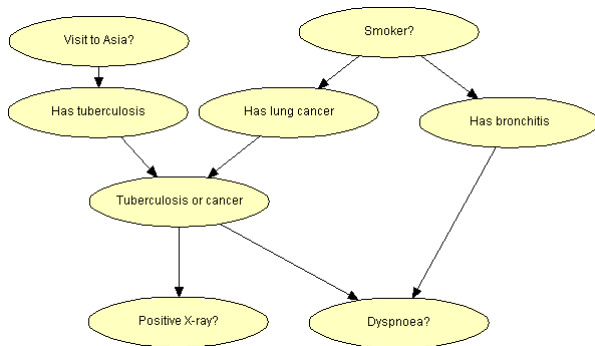
Vyšetřujeme pacienta.

Možná diagnóza je: tuberkulóza, rakovina plic, nebo zánět průdušek.

Zjednodušený diagnostický příklad - plicní klinika

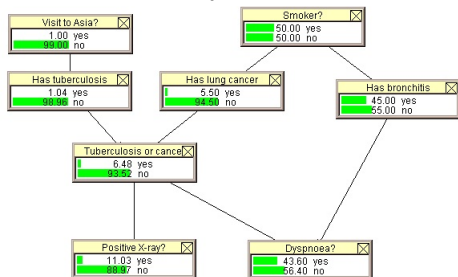
Vyšetřujeme pacienta.

Možná diagnóza je: tuberkulóza, rakovina plic, nebo zánět průdušek.

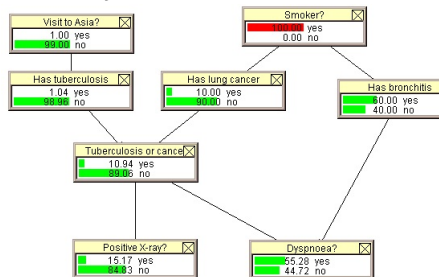


Zjednodušený diagnostický příklad - plicní klinika

Zatím nevíme o pacientovi nic.

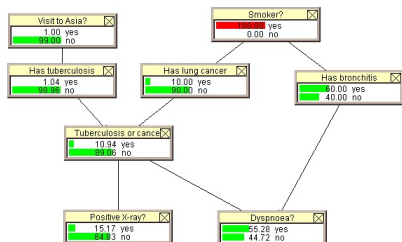


Pacient je kuřák.

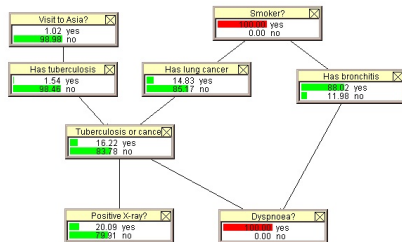


Zjednodušený diagnostický příklad - plicní klinika

Pacient je kuřák

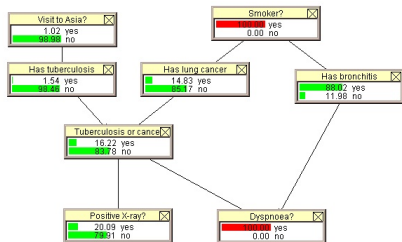


... a je dušný

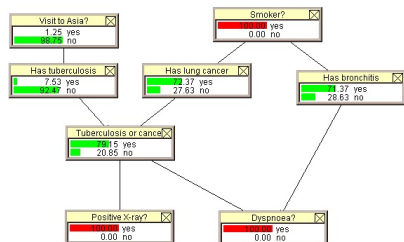


Zjednodušený diagnostický příklad - plicní klinika

Pacient je kuřák, je dušný

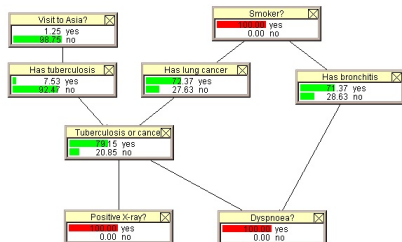


... a navíc jeho RTG je pozitivní

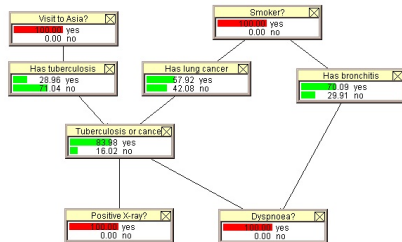


Zjednodušený diagnostický příklad - plicní klinika

Pacient je kuřák, je dušný,
jeho RTG je pozitivní



... a navíc byl v Asii



- Předpokládejme, že máme problém, který budeme modelovat pomocí n veličin a každá veličina může nabývat dvou hodnot.

- Předpokládejme, že máme problém, který budeme modelovat pomocí n veličin a každá veličina může nabývat dvou hodnot.
- Použijeme-li reprezentaci pomocí jedné tabulky potřebujeme pro uložení v paměti počítače distribuce $2^n - 1$ hodnot.

Výhoda reprezentace bayesovskou sítí

- Předpokládejme, že máme problém, který budeme modelovat pomocí n veličin a každá veličina může nabývat dvou hodnot.
- Použijeme-li reprezentaci pomocí jedné tabulky potřebujeme pro uložení v paměti počítače distribuce $2^n - 1$ hodnot.
- Předpokládejme, bayesovskou sítí mající též n veličin nabývajících dvou hodnot s grafem následující struktury:



Výhoda reprezentace bayesovskou sítí

- Předpokládejme, že máme problém, který budeme modelovat pomocí n veličin a každá veličina může nabývat dvou hodnot.
- Použijeme-li reprezentaci pomocí jedné tabulky potřebujeme pro uložení v paměti počítače distribuce $2^n - 1$ hodnot.
- Předpokládejme, bayesovskou sítí mající též n veličin nabývajících dvou hodnot s grafem následující struktury:



- Pro její uložení v paměti počítače potřebujeme $1 + (n - 1) \cdot 2 = 2n - 1$ hodnot.

Výhoda reprezentace bayesovskou sítí

n	$2^n - 1$	$2n - 1$
2	3	3
3	7	5
4	15	7
10	1023	19
100	$1.27 \cdot 10^{30}$	199
1000	$1.07 \cdot 10^{301}$	1999

Typické použití bayesovských sítí

- pro **modelování** a **vysvětlení** chování, problémů v různých oblastech - např. **model chování vody v krajině**,

Typické použití bayesovských sítí

- pro **modelování** a **vysvětlení** chování, problémů v různých oblastech - např. **model chování vody v krajině**,
- pro **přepočtení pravděpodobností** hodnot určitých veličin když jsme pozorovali některé jiné veličiny, t.j. počítání podmíněných pravděpodobností, např. $P(X_{23}|X_{17} = \text{yes}, X_{54} = \text{no})$ - např. **určení diagnózy při pozorovaných příznacích**.

Typické použití bayesovských sítí

- pro **modelování** a **vysvětlení** chování, problémů v různých oblastech - např. **model chování vody v krajině**,
- pro **přepočtení pravděpodobností** hodnot určitých veličin když jsme pozorovali některé jiné veličiny, t.j. počítání podmíněných pravděpodobností, např. $P(X_{23}|X_{17} = \text{yes}, X_{54} = \text{no})$ - např. **určení diagnózy při pozorovaných příznacích**.
- pro nalezení **nejpravděpodobnějších konfigurací** proměnných - např. **při automatickém rozpoznávání řeči**, nebo **při dekódování zakódovaných zpráv**,

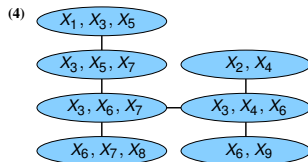
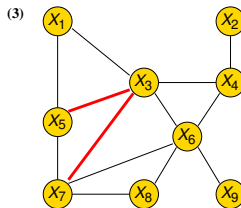
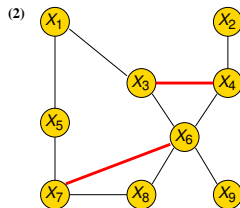
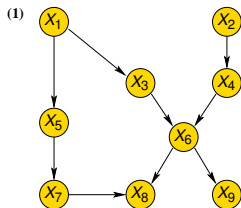
Typické použití bayesovských sítí

- pro **modelování** a **vysvětlení** chování, problémů v různých oblastech - např. **model chování vody v krajině**,
- pro **přepočtení pravděpodobností** hodnot určitých veličin když jsme pozorovali některé jiné veličiny, t.j. počítání podmíněných pravděpodobností, např. $P(X_{23}|X_{17} = \text{yes}, X_{54} = \text{no})$ - např. **určení diagnózy při pozorovaných příznacích**.
- pro nalezení **nejpravděpodobnějších konfigurací** proměných - např. **při automatickém rozpoznávání řeči**, nebo **při dekódování zakódovaných zpráv**,
- pro podporu **rozhodování** při nejisté informaci - použití **teorie maximalizace očekávaného užitku**,

Typické použití bayesovských sítí

- pro **modelování** a **vysvětlení** chování, problémů v různých oblastech - např. **model chování vody v krajině**,
- pro **přepočtení pravděpodobností** hodnot určitých veličin když jsme pozorovali některé jiné veličiny, t.j. počítání podmíněných pravděpodobností, např. $P(X_{23}|X_{17} = \text{yes}, X_{54} = \text{no})$ - např. **určení diagnózy při pozorovaných příznacích**.
- pro nalezení **nejpravděpodobnějších konfigurací** proměných - např. **při automatickém rozpoznávání řeči**, nebo **při dekódování zakódovaných zpráv**,
- pro podporu **rozhodování** při nejisté informaci - použití **teorie maximalizace očekávaného užitku**,
- pro nalezení dobrých **strategií** pro řešení problémů v oblastech z nejistotou - např. **technická diagnostika laserových tiskáren**, nebo **návrh adaptivních testů**.

Výpočty ve stromech spojení



Bayesovské sítě patří mezi pravděpodobnostní grafické modely.

Bayesovské sítě patří mezi pravděpodobnostní grafické modely.
První použití modelů podobných grafickým modelům

- ve statistické mechanice (Gibbs, 1902),
- genetice (Wright, 1921),
- analýze kontingenčních tabulek (Barlett, 1935).

Bayesovské sítě patří mezi pravděpodobnostní grafické modely. První použití modelů podobných grafickým modelům

- ve statistické mechanice (Gibbs, 1902),
- genetice (Wright, 1921),
- analýze kontingenčních tabulek (Barlett, 1935).

Skutečný rozvoj teorie grafických modelů až v osmdesátých letech 20.století:

- Pearl (1982),
- Spiegelhalter a Kill-Jones (1984),
- Perez a Jiroušek (1985),
- Lauritzen a Spiegelhalter (1988)

Vychází základní monografie o grafických modelech a bayesovských sítích:

- Pearl (1988), Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference,
- Neapolitan (1990), Probabilistic reasoning in expert systems: theory and algorithms,
- Hájek, Havránek, Jiroušek (1991), Uncertain Information Processing in Expert Systems,
- Lauritzen (1996), Graphical models,
- Jensen (1996), An introduction to Bayesian networks.

Objevují se aplikace bayesovských sítí v různých oblastech:

- Munin (1989) - Bayesian network for the median nerve
- Pathfinder (1990) - lymph-node pathology,
- Decision theoretic troubleshooting (1994) - technická diagnostika zařízení.

Objevují se aplikace bayesovských sítí v různých oblastech:

- Munin (1989) - Bayesian network for the median nerve
- Pathfinder (1990) - lymph-node pathology,
- Decision theoretic troubleshooting (1994) - technická diagnostika zařízení.

Objevují se první verze software umožňující práci s bayesovskými sítěmi:

- Hugin (1989) <http://www.hugin.com>,
- Netica (1995) <http://www.norsys.com>,