Úvod do bayesovských sítí

Jiří Vomlel

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

30. října 2008

Podmíněná pravděpodobnost

- 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(Horečka=ano)*, nebo jevu *Horečka=ano a Angína=ano*,
  označíme *P(Horečka=ano, Angína=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(Horečka=ano)*, nebo jevu *Horečka=ano a Angína=ano*,

  označíme *P(Horečka=ano, Angína=ano)*
- Podmíněná pravděpodobnost jevu *Horečka=ano* při pozorovaném jevu *Angína=ano*, kterou označíme *P(Horečka=ano | Angína=ano)*, je pravděpodobnost splňující vztah

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

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(Horečka=ano)*, nebo jevu *Horečka=ano a Angína=ano*,

  označíme *P(Horečka=ano, Angína=ano)*
- Podmíněná pravděpodobnost jevu *Horečka=ano* při pozorovaném jevu *Angína=ano*, kterou označíme *P(Horečka=ano* | *Angína=ano)*, je pravděpodobnost splňující vztah

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

- Jestliže *P(Angína=ano)* je nenulová, pak

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

Podmíněná pravděpodobnost - příklad

- Pozorováním jsme získali následující pravděpodobnosti jevů:

$$
\begin{aligned}
P(\mathit{Horečka}=\mathit{ano}, \mathit{Angína}=\mathit{ano}) &= 0.015 \\
P(\mathit{Horečka}=\mathit{ne}, \mathit{Angína}=\mathit{ano}) &= 0.005 \\
P(\mathit{Horečka}=\mathit{ano}, \mathit{Angína}=\mathit{ne}) &= 0.08 \\
P(\mathit{Horečka}=\mathit{ne}, \mathit{Angína}=\mathit{ne}) &= 0.90
\end{aligned}
$$

Podmíněná pravděpodobnost - příklad

- Pozorováním jsme získali následující pravděpodobnosti jevů:

$$\begin{aligned} \mathit{P(Horečka{=}ano,\ Angína{=}ano)} &= 0.015 \\ \mathit{P(Horečka{=}ne,\ Angína{=}ano)} &= 0.005 \\ \mathit{P(Horečka{=}ano,\ Angína{=}ne)} &= 0.08 \\ \mathit{P(Horečka{=}ne,\ Angína{=}ne)} &= 0.90 \end{aligned}$$

- Můžeme spočítat např.

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

Podmíněná pravděpodobnost - příklad

- Pozorováním jsme získali následující pravděpodobnosti jevů:

$$\begin{aligned} \textit{P(Horečka=ano, Angína=ano)} &= 0.015 \\ \textit{P(Horečka=ne, Angína=ano)} &= 0.005 \\ \textit{P(Horečka=ano, Angína=ne)} &= 0.08 \\ \textit{P(Horečka=ne, Angína=ne)} &= 0.90 \end{aligned}$$

- Můžeme spočítat např.

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

- nebo

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

Bayesův vzorec

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

Bayesův vzorec

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

Bayesův vzorec

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

Bayesův vzorec - příklad

- D ... v noci pršelo, 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?

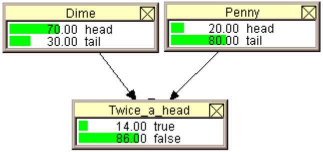
$$
\begin{aligned}
P(D = a|M = a) \\
&= \frac{P(M = a|D = a) \cdot P(D = a)}{P(M = a|D = a) \cdot P(D = a) + P(M = a|D = n) \cdot P(D = n)} \\
&= \frac{\frac{4}{5} \cdot \frac{1}{4}}{\frac{4}{5} \cdot \frac{1}{4} + \frac{2}{5} \cdot \frac{3}{4}} = \frac{\frac{1}{5}}{\frac{1}{5} + \frac{6}{20}} = \frac{\frac{1}{5}}{\frac{1}{2}} = \frac{2}{5}
\end{aligned}
$$

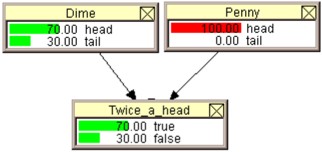
Nezávislost dvou veličin

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

Nezávislost dvou veličin

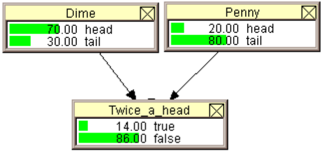
Mějme dvě velmi podivné mince *Dime* and *Penny*. *Dime* má pravděpodobnost, že padne *head* 0.7, ale Penny pouze 0.2.

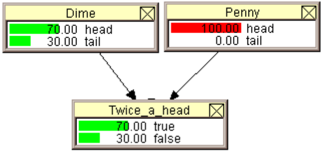




Nezávislost dvou veličin

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



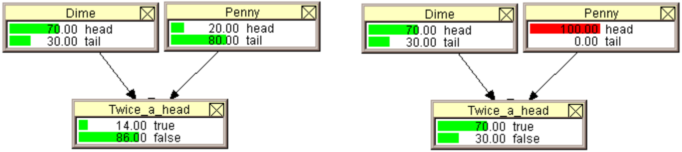


Pravděpodobnost společného výskytu hodnot veličin je rovna součinu pravděpodobností jednotlivých pravděpodobností.

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

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.

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

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

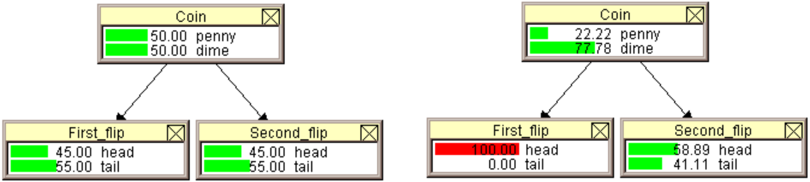
$$P(Dime = head | Penny = head) = P(Dime = head)$$

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.

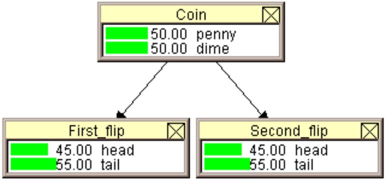
Závislost dvou veličin

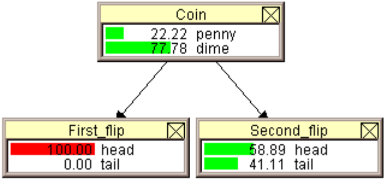
Kolega náhodně vybere jednu z těch dvou podivných mincí a udělá s ní nejprve jeden a pak druhý hod. My nevíme, jakou minci si vybral.



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.





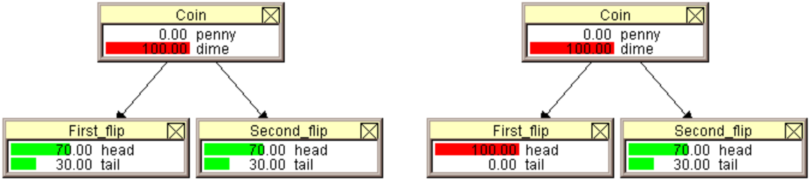
Výsledek prvního hodu má vliv na pravděpodobnost výsledku druhého hodu.

Podmíněná nezávislost dvou veličin

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

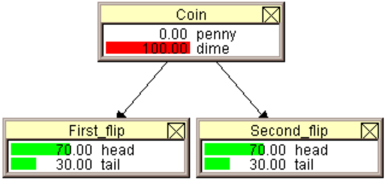
Podmíněná nezávislost dvou veličin

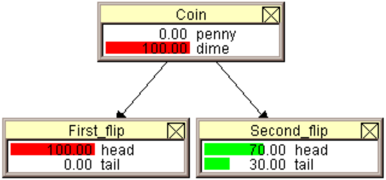
Nyní předpokládejme, že víme, kterou minci si kolega vybral.



Podmíněná nezávislost dvou veličin

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





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

Podmíněná nezávislost

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

$$
\begin{aligned}
& P(\mathit{First\_flip} = \mathit{head}, \mathit{Second\_flip} = \mathit{head} | \mathit{Coin} = \mathit{dime}) \\
= \quad & P(\mathit{First\_flip} = \mathit{head} | \mathit{Coin} = \mathit{dime}) \\
& \cdot P(\mathit{Second\_flip} = \mathit{head} | \mathit{Coin} = \mathit{dime})
\end{aligned}
$$

Podmíněná nezávislost

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

$$\begin{aligned} & P(First\_flip = head, Second\_flip = head | Coin = dime) \\ = \quad & P(First\_flip = head | Coin = dime) \\ & \cdot P(Second\_flip = head | Coin = dime) \end{aligned}$$

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

$$\begin{aligned} & P(Second\_flip = head | Coin = dime, First\_flip = head) \\ = \quad & P(Second\_flip = head | Coin = dime) \end{aligned}$$

Řetězcové pravidlo

Z definice podmíněné pravděpodobnosti plyne, že můžeme psát:

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

Řetězcové pravidlo

Z definice podmíněné pravděpodobnosti plyne, že můžeme psát:

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

Řetězcové pravidlo

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:

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

Proč je Holmesův trávník mokrý?

- Máme 4 veličiny:

Holm	Je Holmesův trávník mokrý?
Rn	Pršelo v noci?
Sprnk	Byl Holmesův postřikovač zapnutý?
Wat	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:

Holm	Je Holmesův trávník mokrý?
Rn	Pršelo v noci?
Sprnk	Byl Holmesův postřikovač zapnutý?
Wat	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:

Holm	Je Holmesův trávník mokrý?
Rn	Pršelo v noci?
Sprnk	Byl Holmesův postřikovač zapnutý?
Wat	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:

Holm	Je Holmesův trávník mokrý?
Rn	Pršelo v noci?
Sprnk	Byl Holmesův postřikovač zapnutý?
Wat	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:

Holm	Je Holmesův trávník mokrý?
Rn	Pršelo v noci?
Sprnk	Byl Holmesův postřikovač zapnutý?
Wat	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:

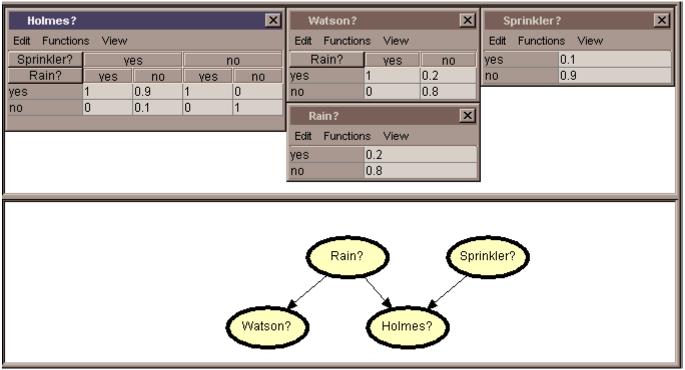
$$
\begin{aligned}
&P(Holm, Wat, Rn, Sprnk) \\
&\quad = P(Holm|Wat, Rn, Sprnk) \cdot P(Wat|Rn, Sprnk) \\
&\quad\quad \cdot P(Rn|Sprnk) \cdot P(Sprnk) \\
&\quad = P(Holm|Rn, Sprnk) \cdot P(Wat|Rn) \cdot P(Rn) \cdot P(Sprnk)
\end{aligned}
$$

Proč je Holmesův trávník mokrý?

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

Proč je Holmesův trávník mokrý?

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



Definice bayesovské sítě

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

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 $\mathbb{X}_i$

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 podmíněné 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 podmíněné 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 podmíněné 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 $P(X_i \mid (X_j)_{j \in pa(i)})$

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

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 $\mathcal{Y}$.

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 $\mathcal{Y}$.

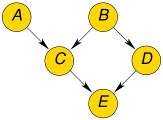
Jestliže i a j jsou d-separovány množinou uzlů $\mathcal{Y}$, pak veličiny X_i a X_j jsou nezávislé dáno $(X_k)_{k \in \mathcal{Y}}$.

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 $(X_k)_{k \in \mathcal{Y}}$.

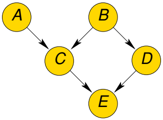


Podmíněné nezávislost dané grafem - d-separace

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 $(X_k)_{k \in \mathcal{Y}}$.



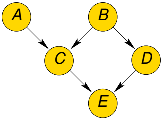
$$A \perp\!\!\!\perp D$$

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 $(X_k)_{k\in\mathcal{Y}}$.



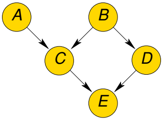
$$A \perp\!\!\!\perp D \mid B$$

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 $(X_k)_{k \in \mathcal{Y}}$.



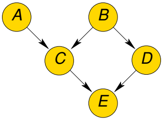
$$A \not\perp\!\!\!\perp D \mid B, E$$

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 $(X_k)_{k \in \mathcal{Y}}$.



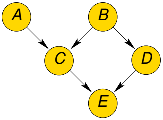
$$C ? D \mid E$$

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 $(X_k)_{k \in \mathcal{Y}}$.



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

Pravděpodobnostní distribuce bayesovské sítě

Použijeme-li řetězcové pravidlo dostaneme:

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

Pravděpodobnostní distribuce bayesovské sítě

Použijeme-li řetězcové pravidlo dostaneme:

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

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:

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

Použijeme-li podmíněné nezávislosti z grafu pak dostaneme

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

Toto je pravděpodobnostní distribuce representovaná bayesovskou sítí.

Pravděpodobnostní distribuce bayesovské sítě

Použijeme-li řetězcové pravidlo dostaneme:

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

Použijeme-li podmíněné nezávislosti z grafu pak dostaneme

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

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:

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

Použijeme-li podmíněné nezávislosti z grafu pak dostaneme

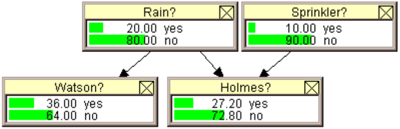
$$P((X_i)_{i \in V}) = \prod_{i \in V} P(X_i \mid (X_j)_{j \in pa(i)})$$

Toto je pravděpodobnostní distribuce representovaná bayesovskou sítí.

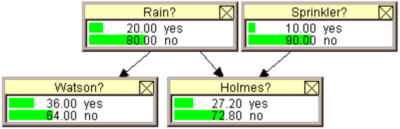
Pro tuto distribuci platí:

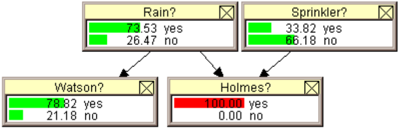
- splňuje podmíněné nezávislosti zakódované v grafu a
- její podmíněné pravděpodobnostní distribuce odpovídají $P(X_i \mid (X_j)_{j \in pa(i)})$, $i \in V$.

Byl Holmesův postřikovač zapnutý?

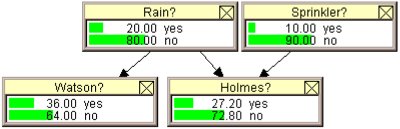


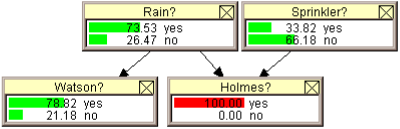
Byl Holmesův postřikovač zapnutý?

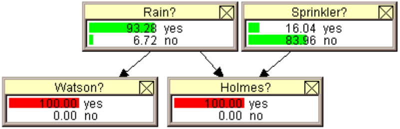




Byl Holmesův postřikovač zapnutý?





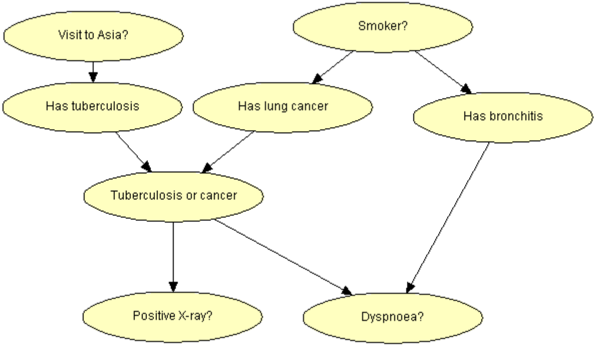


Zjednodušený diagnostický příklad - plicní klinika

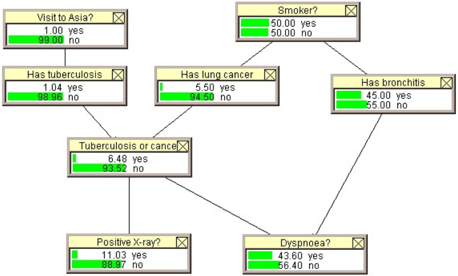
Vyšetřujeme pacienta.
Možná diagnóza je: tuberkulóza, rakovina plic, nebo zánět průdušek.

Zjednodušený diagnostický příklad - plicní klinika

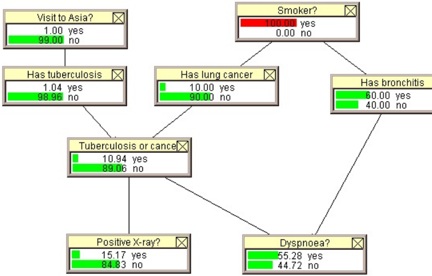
Vyšetřujeme pacienta.
Možná diagnóza je: tuberkulóza, rakovina plic, nebo zánět průdušek.



Zatím nevíme o pacientovi nic.

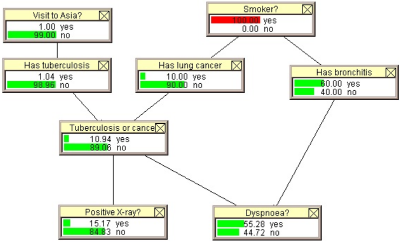


Pacient je kuřák.

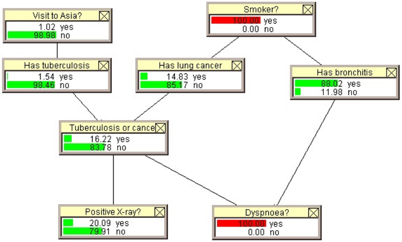


Zjednodušený diagnostický příklad - plicní klinika

Pacient je kuřák

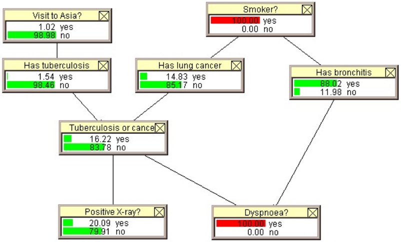


... a je dušný

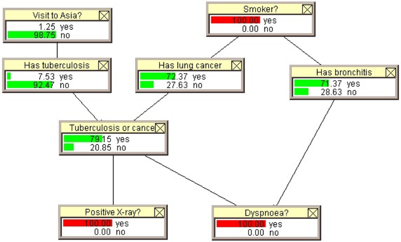


Zjednodušený diagnostický příklad - plicní klinika

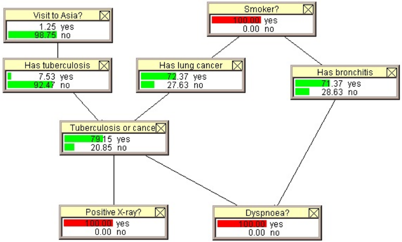
Pacient je kuřák, je dušný



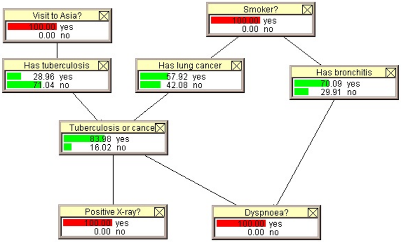
... a navíc jeho RTG je positivní



Pacient je kuřák, je dušný, jeho RTG je positivní



... a navíc byl v Asii



Výhoda reprezentace bayesovskou sítí

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

Výhoda reprezentace bayesovskou sítí

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

$$X_1 \longrightarrow X_2 \longrightarrow X_3 \longrightarrow \cdots\cdots \longrightarrow X_n$$

- 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. $P(X_{23} | X_{17} = yes, X_{54} = no)$ - 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. $P(X_{23}|X_{17} = yes, X_{54} = no)$ - např. určení diagnózy při pozorovaných příznacích.
- pro nalezení nejpravděpodobnějších konfigurací proměných - např. při automatickém rozpoznávání řeči, nebo při dekódování zakódovaných zpráv,

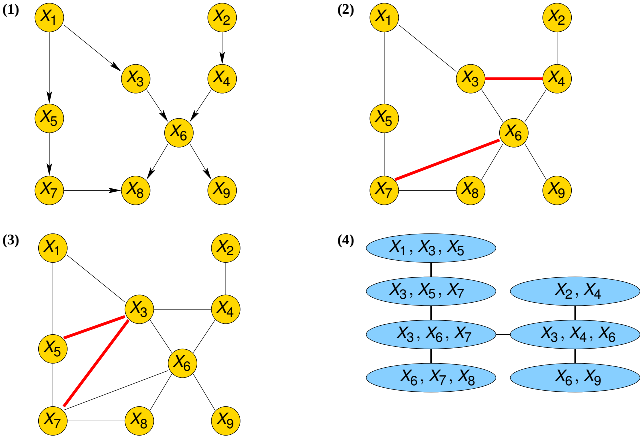
Typické použití bayesovských sítí

- 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. $P(X_{23}|X_{17} = yes, X_{54} = no)$ - 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. $P(X_{23}|X_{17} = yes, X_{54} = no)$ - 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



Z historie bayesovských sítí - 1

Bayesovské sítě patří mezi pravděpodobnostní grafické modely.

Z historie bayesovských sítí - 1

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).

Z historie bayesovských sítí - 1

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 torie grafických modelů až v osmdesátých letech 20.století:

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

Z historie bayesovských sítí - 2

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.

Z historie bayesovských sítí - 3

Objevují se aplikace bayesovských sítích v různých oblastech:

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

Z historie bayesovských sítí - 3

Objevují se aplikace bayesovských sítích 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`,