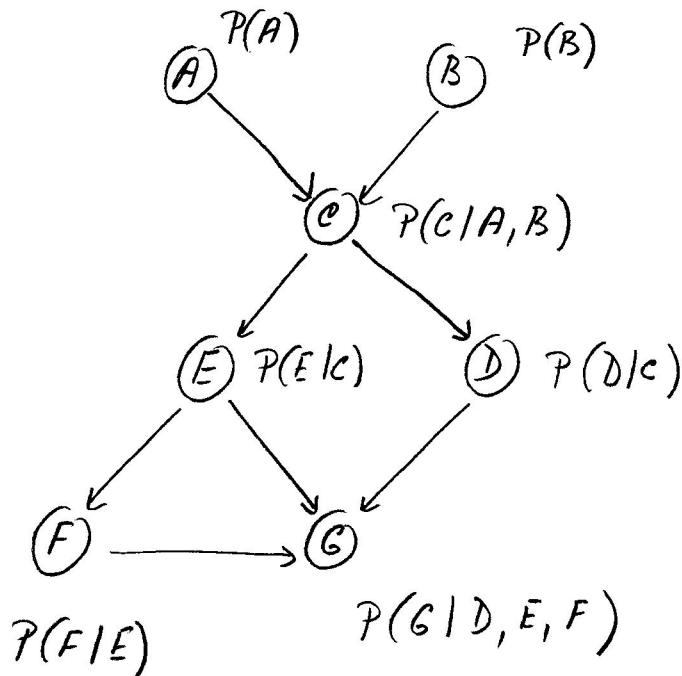


Zatím jsme si malovali sítě jen jako propojení proměnné, tak teď ještě jednu s příslušnými pravděpodobnostmi:



Počítání pravděpodobností v Bayesovské síti

Ornacime $U = \{A_1, \dots, A_n\}$ prostor všech proměnných v Bayesovské síti

$$P(U) = P(A_1, \dots, A_n) \text{ jejich sdružené rozdelení}$$

Pokud by A_1, \dots, A_n byly nezávislé, pak by platilo

$$P(A_1, \dots, A_n) = \prod_{i=1}^n P(A_i)$$

Pokud by byly všechny závislé, pak by naopak jejich sdružené rozdelení bylo prakticky nerečitelné (protože z veličin A_i nabyvá konečně mnoha hodnot a sdružené rozdelení musí popsat všechny možné jejich kombinace)

V Bayesovské sítí ale můžeme počítat jednodušeji podle vzorce

$$P(U) = \prod_i P(A_i | \text{rodice}(A_i)) \quad (\text{tzv. "řetězové pravidlo")}$$

Je to podobné jako v Markovových řetězcích, kde počítáme pravděpodobnost posloupnosti výsledků pokud jde o

$$P(E_0, E_1, \dots, E_m) = P(E_0)P(E_1|E_0) \dots P(E_m|E_{m-1})$$

(Markovský proces by se dal zahájet jako jednoduchá B.S. se sériovým propojením)

Příklad s autem

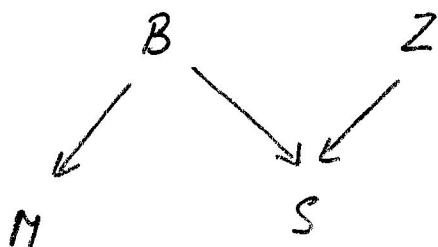
označime B... benzín (ano, ne)

M... měřicí paliva (plno, 1/2, prázdro)

Z... zapalování (čisté, zanesené)

S ... start (ano, ne)

propojení:



$$P(B, M, Z, S) = P(B)P(Z)P(M|B)P(S|M, Z)$$

Budeme předpokládat, že auto je teoreticky v porování (tj. množství benzínu a nezanesené zapolování) a vyjdeme z počátečních pravděpodobnosti $P(B) = (0.98, 0.02)$ a $P(Z) = (0.96, 0.04)$

} vektor

Podmíněné pravděpodobnosti $P(M|B)$ budou odvozeny fakt, že ukazatel paliva může ukazovat malinko nepřesně a že nejpravděpodobnější je, že rádce je přibližně v polovině plna' máme tabulku:

	$B = \text{ano}$	$B = \text{ne}$	
$M = \text{plno}$	0.39	0.001	dvoourázové
$M = 1/2$	0.60	0.001	tabulka
$M = \text{prázdro$	0.01	0.998	

Podmíněné pravděpodobnosti $P(S|B, Z)$:

<u>$S = \text{ano}$</u>		<u>$S = \text{ne}$</u>	
	$B = \text{ano}$	$B = \text{ne}$	
$Z = \text{čisté'}$	0.99	0	0.01 1
$Z = \text{zanesené'}$	0.01	0	0.99 1

trojrozměrná tabulka

Sdružené rozdělení $P(B, M, Z, S)$ už vyžaduje 4-rozměrnou tabulku, proto ho napišeme jako 2 trojrozměrné, zvlášt pro $S = \text{ano}$ a $S = \text{ne}$

$S = \text{ano}$

	$M = \text{plno}$	$M = 1/2$	$M = \text{prázdro}$
$Z = \text{čisté'}$	(0.363, 0)	(0.559, 0)	(0.0093, 0)
$Z = \text{zanesene'}$	(0.00015, 0)	(0.00026, 0)	($3.3 \cdot 10^{-6}$, 0)

$S = \text{ne}$

	$M = \text{plno}$	$M = 1/2$	$M = \text{prázdro}$
$Z = \text{čisté'}$	(0.00367, $1.9 \cdot 10^{-5}$)	(0.00564, $1.9 \cdot 10^{-5}$)	($9.4 \cdot 10^{-5}$, 0.0192)
$Z = \text{zanesene'}$	(0.01514, $8 \cdot 10^{-4}$)	(0.0233, $8 \cdot 10^{-4}$)	(0.000388, 0.000798)

čísla v závorkách jsou pro $B = \text{ano}$, $B = \text{ne}$

Vypočt např. $P(B = \text{ano}, M = 1/2, Z = \text{čisté'}, S = \text{ne}) =$

$$= P(B = \text{ano}) P(Z = \text{čisté'}) P(M = 1/2 | B = \text{ano}) \cdot P(S = \text{ne} | B = \text{ano}, Z = \text{čisté'}) =$$

$$= 0.98 \cdot 0.96 \cdot 0.6 \cdot 0.01 = 0.0056448 = 0.00564$$

Náš úloha byla: zjistili jsme, že auto nenastartovalo,
čehož vědět, cím to bylo

tj. počítáme $P(Z|S=\text{ne})$, $P(B|S=\text{ne})$

$$P(Z|S=\text{ne}) = \frac{P(Z, S=\text{ne})}{P(S=\text{ne})}$$

$$P(Z, S=\text{ne}) = \sum_B \sum_M P(B, M, Z, S=\text{ne})$$

dostaneme výsčítáním řádku z tabulky s dřízeným
rozdělením pro $S=\text{ne}$

$$\begin{aligned} P(Z = \text{ciste}, S=\text{ne}) &= 0.00367 + 0.00564 + 9.6 \cdot 10^{-5} + \\ &+ 1.9 \cdot 10^{-5} + 1.9 \cdot 10^{-5} + 0.0192 = \\ &= 0.028642 \end{aligned}$$

$$P(Z = \text{zanesene}, S=\text{ne}) = 0.03965$$

$P(S=\text{ne})$ dostaneme dalším výsčítáním ještě přes 2

$$P(S=\text{ne}) = \sum_Z P(Z, S=\text{ne}) = 0.068292$$

$$P(Z = \text{ciste} | S=\text{ne}) = \frac{0.028642}{0.068292} \doteq 0,42$$

$$P(Z = \text{zanesene} | S=\text{ne}) \doteq 0.58$$

Podobně dostaneme $P(B=ano | S=ne) =$

$$= \frac{\sum_{\mathcal{Z}} \sum_M P(B=ano, \mathcal{Z}, M, S=ne)}{P(S=ne)} = 0.71$$

$$P(B=ne | S=ne) = 0.29$$

Po z toho plyně?

Vypadá to, že príčina je v záplatování, ale není to príčinou přesředcív!

Můžeme to zprávnit - když jsme nevystartovali, pravděpodobně jsme se podívali na ukazatel paliva. Nechť např. ukazoval $\frac{1}{2}$.

Procedeme tedy výpočty, ale z tabulky sdruženého rozdělení pro $S=ne$ použijeme jen sloupec $M=\frac{1}{2}$, tj.

$$P(B, \mathcal{Z}, M=\frac{1}{2}, S=ne)$$

	$B=ano$	$B=ne$
$\mathcal{Z}=čisté'$	0.00564	$1.9 \cdot 10^{-5}$
$\mathcal{Z}=zanesene'$	0.0233	$8 \cdot 10^{-7}$

V tomto případě nám vypadá:

$$P(B=ano | S=ne, M=\frac{1}{2}) = 0.999$$

$$P(B=ne | S=ne, M=\frac{1}{2}) = 0.001$$

$$P(\mathcal{Z}=čisté' | S=ne, M=\frac{1}{2}) = 0.196$$

$$P(\mathcal{Z}=zanesene' | S=ne, M=\frac{1}{2}) = 0.804$$

z příkladu je vidět, že i s použitím řetězového pravidla je výpočet komplikovaný, da'se ale zjednodušit.

Např. když víme, že auto nenastartovalo, je zbytečné počítat tabulku sdruženého rozdělení pro $S = \text{ano}$

když navíc víme, že ukazatel paliva ukazoval $\frac{1}{2}$, stačí nám pouze sloupec pro $M = \frac{1}{2}$.

Máme-li tedy evidenci $e = (S = \text{ne}, M = \frac{1}{2})$, počítáme

$$\begin{aligned} P(U, e) &= P(B, \chi, M = \frac{1}{2}, S = \text{ne}) = \\ &P(B) P(\chi) P(M = \frac{1}{2} | B) P(S = \text{ne} | B, \chi) \end{aligned}$$

Když navíc počítáme $P(B | e) = P(B | M = \frac{1}{2}, S = \text{ne})$,

pak marginalizujeme: (tj. vyeliminujeme proměnnou χ)

$$\begin{aligned} P(B, M = \frac{1}{2}, S = \text{ne}) &= \sum_{\chi} P(B, \chi, M = \frac{1}{2}, S = \text{ne}) = \\ &= P(B) \sum_{\chi} P(\chi) P(M = \frac{1}{2} | B) P(S = \text{ne} | B, \chi) = \\ &= P(B) P(M = \frac{1}{2} | B) \sum_{\chi} P(\chi) P(S = \text{ne} | B, \chi) \end{aligned}$$

Tímto způsobem se da' snížit prostorová' náročnost výpočtu -

nemusím držet v paměti všechny tabulky 'rozdělení' - po marginalizaci přes χ můžu $P(\chi)$ a $P(S = \text{ne} | B, \chi)$ zapomenout a nechat si pouze výsledek