

Bayesovské sítě

①

Je to způsob grafického vyjádření závislosti mezi jedy (proměnnými),
mírou závislosti je podmíněná pravděpodobnost.

Slouží k rozhodování v situacích, které jsou zatíženy nějakou
neurčitostí.

Jednoduché příklady

- 1) Máš auto a 4 proměnné: benzín (ano, ne)
ukazatel paliva (prázdná nádrž,
1/2, plná)
zapalování (čisté, zanesené)
start (ano, ne)

Vím, že např.:

kdýž nebudu mít benzín, pravděpodobně nenastartuju

-"-

ukazatel pravděpodobně ukáže prázdnou

kdýž je zanesené zapalování, pravděpodobně nenastartuju

kdýž mám dost paliva a čisté zapalování, pravděpodobně
nastartuju

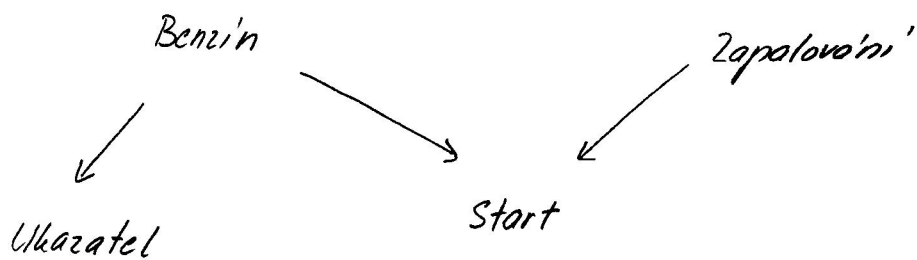
Můj problém ale bude přesně opačný - zjistila jsem, že
auto nestartuje a odhaduji, čím to pravděpodobně bude

Podívám se na ukazatel paliva, zjistím, že ukazuje 1/2 -

můžu z toho usoudit, že problém je v zapalování?

nebo ukazatel ukazuje špatně a je možné, že nemám benzín?

graficky :

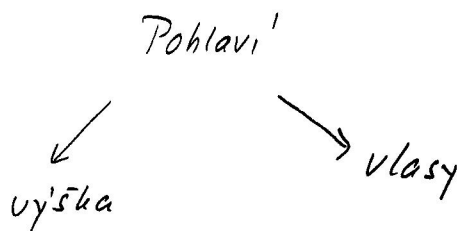


- 2) proměnné : pohlaví (muž, žena)
 výška (malá, velká)
 délka vlasů (krátké, dlouhé)

Vím, že : muž bude pravděpodobně vyšší
 ——— mít kratší vlasy

Můj problém : vidím osobu s dlouhými vlasy
 bude to spíš muž nebo žena ?
 bude spíš menší nebo vyšší ?

graficky :



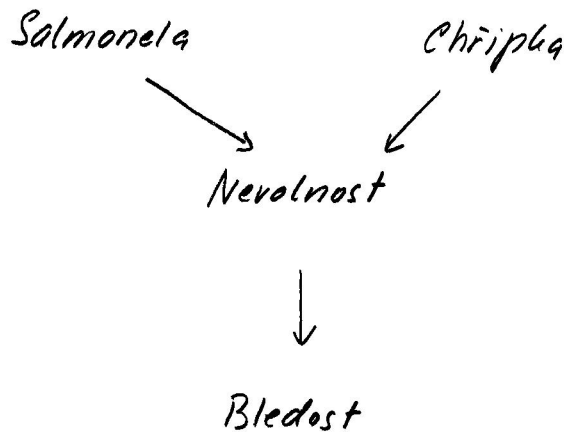
3) proměnné : salmonelóza (ano, ne)

chřipka (ano, ne)

žaludeční nevolnost (ano, ne)

bledost (ano, ne)

graf:



Bayesovská síť je orientovaný acyklický graf

vrcholy jsou náhodné proměnné s konečnou množinou navzájem disjunktivními stavy

každé proměnné A , jejímiž rodiči jsou $B_1, \dots, B_n$, je přiřazena podmíněná pravděpodobnost $P(A|B_1, \dots, B_n)$

Podmíněné pravděpodobnosti

Definice: $P(A|B) = \frac{P(AB)}{P(B)}$ pro $P(B) > 0$

Obecnější definice: $P(A|B, C) = \frac{P(AB|C)}{P(B|C)}$

aplikuje se v případech, kdy všechny jevy uvažujeme jako podmíněné nějakým kontextem C

(Reyniho axiomatika podmíněných pravděpodobností)

Z definic dále plyne: (Bayesův vzorec)

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$$

$$\text{resp. } P(B|A, C) = \frac{P(A|B, C)P(B|C)}{P(A|C)}$$

Jevy A, B jsou nezávislé, když $P(AB) = P(A)P(B)$

$$\text{neboli } P(A|B) = P(A)$$

Jevy A, B jsou podmíněně nezávislé při daném C , když

$$P(A|C) = P(A|BC)$$

Dále platí:

Věta o úplné pravděpodobnosti: Necht' $B_1, \dots, B_m$ jsou jevy takové, že $P(B_i) > 0$ pro vs. i , $B_i B_j = \emptyset$ pro vs. $i \neq j$, $\sum_{i=1}^m P(B_i) = 1$.

$$\text{Pak } P(A) = \sum_{i=1}^m P(AB_i) = \sum_{i=1}^m P(A|B_i)P(B_i).$$

Bayesova věta:

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^m P(A|B_i)P(B_i)}$$

Terminologie: $P(B_k)$... apriorní pravděpodobnost jevu B_k

$P(B_k|A)$... posteriorní —

Věta o násobení pravděpodobnosti: Necht' $P(A_1, \dots, A_{n-1}) > 0$.

$$\text{Pak } P(A_1, \dots, A_n) = P(A_1)P(A_2|A_1)P(A_3|A_2A_1) \dots P(A_n|A_1 \dots A_{n-1})$$

$$\text{Dě: } P(A_1) \cdot \frac{P(A_2 A_1)}{P(A_1)} \cdot \frac{P(A_3 A_2 A_1)}{P(A_2 A_1)} \dots \frac{P(A_n \dots A_1)}{P(A_{n-1} \dots A_1)}$$

Pr: V Morseově abecedě jsou znaky "tečka" a "čárka" v poměru 5:3.

Přenos zpráv má tyto statistické vlastnosti: u vyslaných signálů

"tečka" přijde $\frac{4}{5}$ jako "čárka" a u přijatých signálů "čárka"

přijde $\frac{1}{3}$ jako "tečka". Přijali jsme signál "tečka". Jaka je pst., že byl skutečně vyslán signál "tečka"?

Ře: $A_1 \dots$ byl vyslán signál "tečka" $P(A_1) = \frac{5}{8}$

$A_2 \dots$ — " — "čárka" $P(A_2) = \frac{3}{8}$

$B \dots$ byl přijat signál "tečka"

$$P(B|A_1) = \frac{3}{5} \quad P(B|A_2) = \frac{1}{3}$$

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} = \frac{\frac{3}{5} \cdot \frac{5}{8}}{\frac{3}{5} \cdot \frac{5}{8} + \frac{1}{3} \cdot \frac{3}{8}} =$$

$$= \frac{\frac{3}{8}}{\frac{4}{8}} = \frac{3}{4} = 0,75$$

Kabaxnik vybra obraz ze skupiny 10 originalu a 2 kopie. Radi' u soborniku,

kluj' roxena' original a 2 kopie s pravdepodobnosti $\frac{5}{6}$. Atornik vrati, ke vybrany obraz je original. Jaka je pravdepodobnost toho, ze to je skutocne original?

roz. A. obraz je original

B. obraz je kopia

C. atornik selne, ze obraz je original

$$P(A) = \frac{10}{12}$$

$$P(B) = \frac{2}{12}$$

$$P(C|A) = \frac{5}{6} \quad P(C|B) = \frac{1}{6}$$

$$P(A|C) = \frac{P(C|A)P(A)}{P(C|A)P(A) + P(C|B)P(B)} = \frac{\frac{5}{6} \cdot \frac{10}{12}}{\frac{5}{6} \cdot \frac{10}{12} + \frac{1}{6} \cdot \frac{2}{12}} = \frac{5 \cdot 10}{5 \cdot 10 + 2} = \frac{50}{52} = 0,96$$

V Bayesovských sítích se kombinují 3 základní typy propojení:

1) sériové

$$A \longrightarrow B \longrightarrow C$$

A ovlivňuje B, B ovlivňuje C, A ovlivňuje C prostřednictvím B

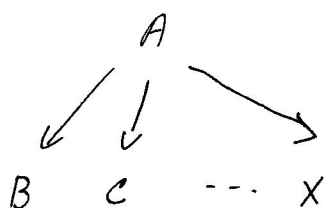
to platí, když máme možnost přímo pozorovat jevy A nebo C
jakmile je známo B, tak jev A už neovlivní C

(jevy A a C budou ^{podmíněně} nezávislé při daném B)

Terminologie: jevy A a C jsou d-separované při daném B %
(jinak také strukturnálně nezávislé)

nejdou-li d-separované, jsou d-propojené (nemusí mezi nimi být přímo hrana)

2) divergentní



A může mít vliv na všechny jevy B, C, ..., X

jakmile známe A, jevy B, C, ..., X se stanou ^{podmíněně} nezávislými

(jsou d-separované při pevném A)

příklad: pohlaví nějakým způsobem má vliv na délku vlasů

— " —

výška postavy

ale délka vlasů a výška postavy u mužů jsou nezávislé

(není pravda, že by třeba vyšší muži měli delší vlasy)

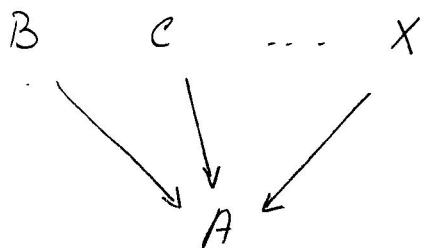
Příklad: A... množství benzínu v nádrži auta včera
B... — " — dnes
C... — " — zítra

S autem se jezdí, ale palivo se doplňuje, tj. může jediné
ubývat.

Vím-li, že dnes ^{skoro} je nádrž plná, pak to, kolik bude
benzínu zítra, už nezávisí na tom, kolik ho bylo včera

3) konvergentní

(7)



Tady je to obráceně - dokud není nic známo o A, jsou jevy

$B, C, \dots, X$ nezávislé

příklad: dokud nevím, jestli má člověk žaludeční potíže nebo ne, tak informace o tom, že nemá salmonelózu, nic neříká o možnosti, že má chřipku

Jakmile ale vím, že má žaludeční potíže a nemá salmonelózu, tak pravděpodobnost chřipky roste.

Totož ale bude platit, když místo nevolnosti budu evidovat bledost.

Separující v tomto případě nemusí být přímo jev A, ale i jeho následník.

Pojem evidence - je to údaj o určitosti stavů náhodně proměnné

"hard evidence" - tu mám v případě, že proměnnou přímo pozoruji a její stav zjistím (např. že pozorovaná osoba je muž)

"soft evidence" - ostatní případy

Další pojmy: certainty - jistota, určitost
belief - víra

(ono se to většinou formálně nedefinuje, spíš intuitivně a z kontextu je zřejmé, o co jde)

Jestliže např. proměnná A může mít stavy $a_1, \dots, a_n$ s pravděpodobnostmi $x_1, \dots, x_n$ a evidence říká, že momentálně může být jen ve stavech a_i nebo a_j , pak

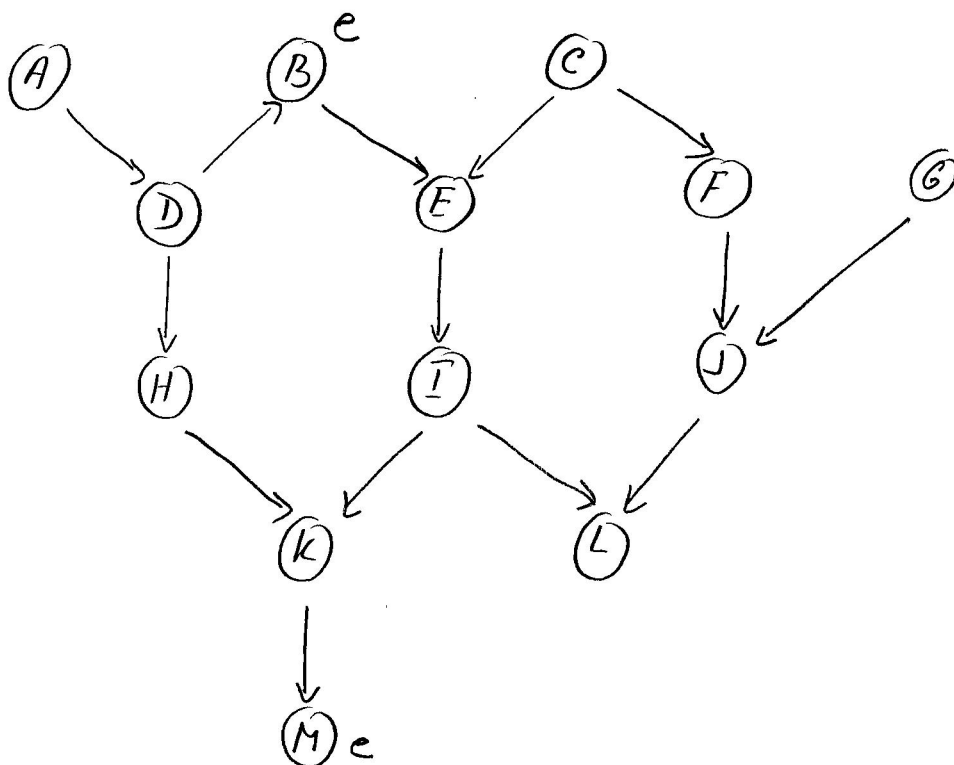
$$P(A) = P(a_1, \dots, a_n) = (x_1, \dots, x_n)$$

$$P(A, e) = (0, \dots, x_i, 0, \dots, x_j, 0, \dots, 0) =$$

$$= P(A) \times \underbrace{(0, \dots, 1, 0, \dots, 1, 0, \dots)}_{\text{finding (nálezy?)}} = P(a)_e$$

Příklady kombinací propojení:

a)



9

e u proměnných B a M značí, že jim je přiřazena evidence
(tj. jejich stavy jsou pozorované, tudíž pevné)

ktelé stavy v síti může ovlivnit proměnná A a které ne
(tj. od kterých proměnných je A d-separována?)

Propojení $A \rightarrow B \rightarrow E$ je sériové a B je pevné, tj.
toto spojení je blokováno a A nemůže ovlivnit E.

Ale: A je propojeno s $D \rightarrow H \rightarrow K$

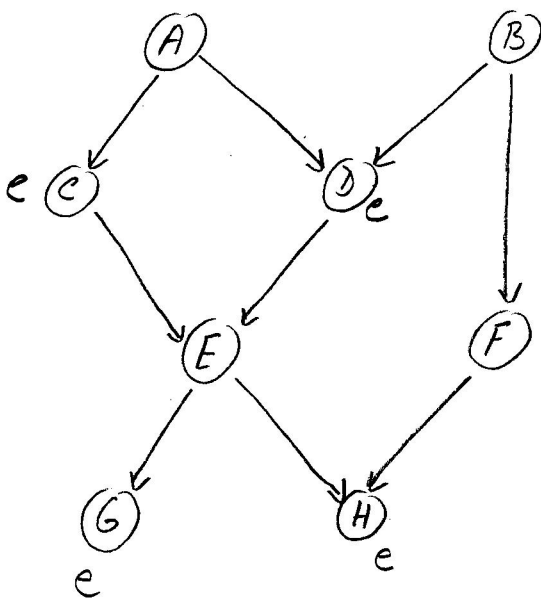
propojení $H - K - I$ je konvergentní a protože

přímý potomek K, tj. M je pevné, pak H je propojeno

s I a tudíž i se zbytkem sítě kromě G

($C - E - I$ je sériové a není blokováno, takže $C - F - J - L$,
ale $F - G - J$ je konvergentní a J není pevné)

b)



Přestože všechny proměnné

v okolí E jsou pevné,

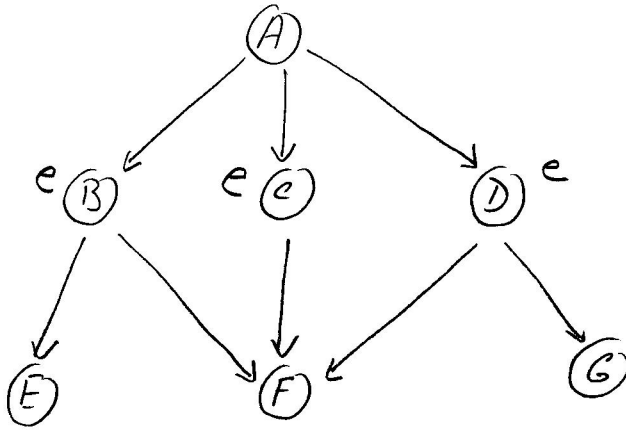
je E propojeno s F

(H je pevné a propojení $E - H - F$
je konvergentní), tudíž i s B

a také s A (D je pevné,

propojení $A - D - B$ je konvergentní)

e)



Tady je F separované od A, E, G

Terminologie: když proměnné A, B nejsou d -separované,
nazývají se d -propojené

d -separované proměnné se někdy nazývají strukturně
nezávislé

Plati: Jsou-li A a B d -separované, pak změna "jistoty"
 A nemá vliv na "jistotu" B .

Jsou-li A a B d -propojené, pak změna "viru" v A
nemusí nutně změnit "viru" v B .