Modelování neurčitosti ve znalostních systémech

Jiří Kléma

Katedra kybernetiky,
FEL, ČVUT v Praze

http://ida.felk.cvut.cz/moodle/

Struktura přednášky

- Neurčitost (nejistota, vágnost)
  - vymezení pojmu, příklady, původ nejistoty.
- Práce s neurčitostí ve znalostních systémech
  - velmi zjednodušené modely,

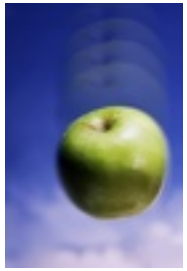
    $\Downarrow$
  - modelování sdruženou pravděpodobností.
- Příklady konkrétních znalostních systémů
  - MYCIN,
  - PROSPECTOR,
  - a srovnání se základním bayesovským přístupem.

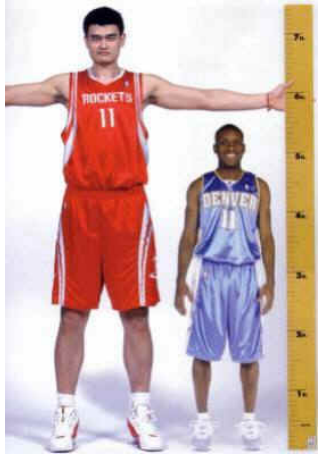
Neurčitost

- omezená znalost, neschopnost přesně popsat existující stav nebo výsledek možných akcí.
- související pojmy
  - **riziko** – nejistota, jejímž možným následkem je ztráta či akce s nežádoucím následkem,
  - **přesnost** – blízkost mezi výsledkem měření a skutečnou měřenou hodnotou,
  - **pravděpodobnost** – slouží ke kvantifikaci nejistoty.
- Původ neurčitosti
  - neúplná pozorovatelnost prostředí
    - lokální pozorovatelnost (viz. wumpus), vysoká cena za získání informace atd.
  - šumy a chyby (měření, senzory),
  - přílišná složitost úplného modelování, nutnost zjednodušení
    - mj. kvalifikační problém – nelze explicitně uvést a kontrolovat všechny podmínky (často jen naprosto výjimečně nesplněné) nutné pro provedení akce.
  - nejisté výsledky akcí
    - nemožnost provést akci identicky, změny v nepozorovatelném prostředí.
  - nedeterministické vztahy
    - ať už skutečně nebo zdánlivě vlivem přílišné složitosti systému,
    - symptomy naznačují onemocnění, to ale není jejich přímým a jediným jejich důsledkem.

Neurčitá znalost - příklady

- Situace s použitelnou přesnou znalostí jsou v reálném světě vzácné
  - příklady několika pravidel (další neurčitost plyne z nejistých dat)
  - *"Pustíme-li předmět, který držíme v ruce, předmět padá dolů."*
    - co když bude předmět lehčí než vzduch?
    - co když budu ve vodě, magnetickém poli, vesmíru?
    - ZÁKON, má ale příliš výjimek pro přesný zápis,
    - kvalifikační problém.
  - *"Pacient se zápalem plic má zvýšenou teplotu."*
    - jde o typický projev, vzácně nemusí platit a není jasné proč,
    - pravidlo běžně používáno v součinnosti s dalšími.
  - *"Všichni vynikající hráči košíkové jsou vysocí."*
    - pravidlo občas neplatí, stejně jako u minulého příkladu,
    - vágnost pojmů vynikající a vysoký,
    - mj. vysoký člověk není vysoký basketbalista,
    - hráči okolo 180cm jsou malí.





Triviální modely pro práci s neurčitostí

- trojhodnotová logika

A	$\neg$ A
0	1
X	X
1	0

A $\wedge$ B	0	X	1
0	0	0	0
X	0	X	X
1	0	X	1

A $\Rightarrow$ B	0	X	1
0	1	1	1
X	X	X	1
1	0	X	1

  - X – neznámá (chybějící) hodnota,
  - vyhodnocení pravidla: konjunkce dílčích podmínek – všechny musí být splněny
    0 je agresivní, 1 je neutrální, X střední agresivita
  - kombinace pravidel se stejnou hypotézou: disjunkce závěrů pravidel
    stačí jediné splněné pravidlo, aby hypotéza platila, 1 je agresivní
  - příklady odvozování:
    stav světa: A=0, B=1, C=X, D=1, E=1
    pravidla: P1: A,B,D $\Rightarrow$ H, P2: C,E $\Rightarrow$ H, P3: B,D,E $\Rightarrow$ H
    výstup: B{P1} $\rightarrow$ H=0, B{P2} $\rightarrow$ H=X, B{P3} $\rightarrow$ H=1, B{P1,P2,P3}$\rightarrow$ H=1

- Použití:
  - databáze – jazyk SQL (NULL), elektronika – logické obvody,
  - v expertních systémech: NExpert Object, KappaPC.

Proč ne (jenom čistá) pravděpodobnost?

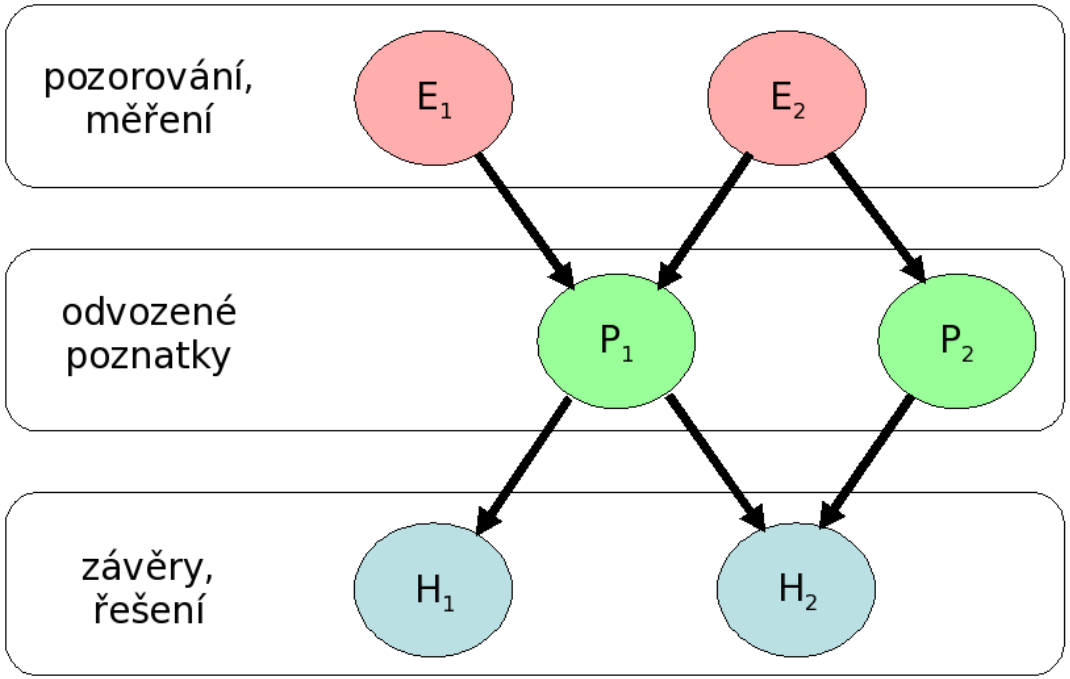
- vychází z relativní četnosti - je objektivní
  - třetina úmrtí je způsobena cévními onemocněními,
- řada tvrzení je subjektivních
  - řekl bych, že cévní onemocnění způsobí čtvrtinu úmrtí,
  - epistemická modalita,
- vágní tvrzení nevycházejí z relativní četnosti
  - Petr je docela vysoký,
  - fuzzy logika,
- tyto námitky částečně zohledňuje bayesovská interpretace psti (vs dříve uvedená frekventistická)
- ale především je pro reálné úlohy nepraktická (nezvladatelná)
  - 2^n pravděpodobností pro určení vztahů s n výroky,
  - nereálné pro experta a stejně tak obtížné experimentální stanovení z dat,
  - i pokud bychom znali, stále exponenciální časové a paměťové nároky
    - zřejmé u sdružené spojité distribuční funkce,
    - **prokletí dimenzionality** - potřeba pozorování roste exp s počtem proměnných.

Práce s neurčitostí (v expertních systémech)

- pseudopravděpodobnostní
  - Prospector, FEL-Expert,
- míry důvěry
  - Mycin,
- algebraická teorie
  - Equant,
- nemonotonní logika
  - Prolog,
- fuzzy logika
  - Sak, Nest,
- pravděpodobnostní (příště)
  - bayesovské pravidlo a sítě.

Diagnostický systém

- znalost ve formě acyklického orientovaného grafu,
- rozlišuje dotazovatelné, vnitřní a cílové uzly,
  - (omezeně) pozorovatelné předpoklady, pomocné proměnné a nepozorovatelné závěry,
  - hledáme rozumné závěry pro danou množinu pozorování,
  - zde výhradně nejistota, stranou mj. řídicí nebo vysvětlovací mechanismus.

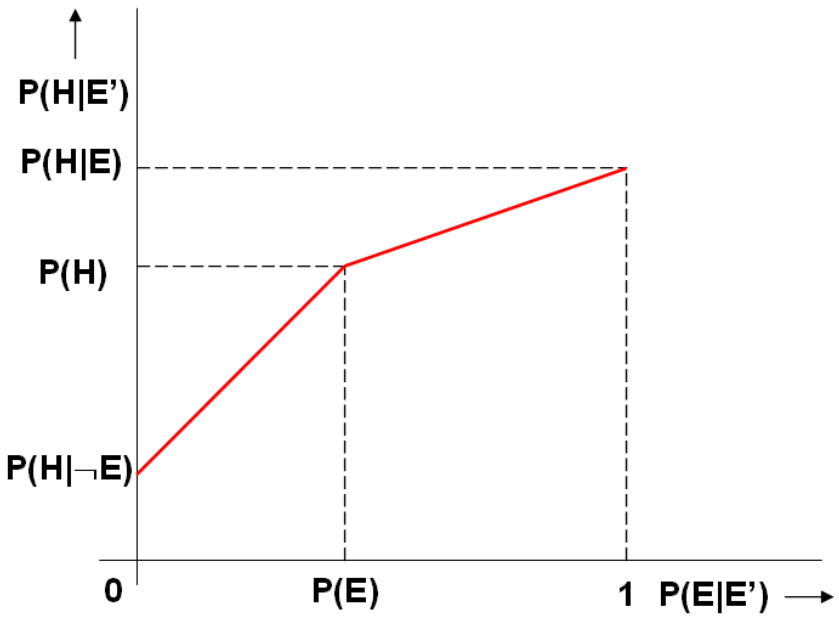


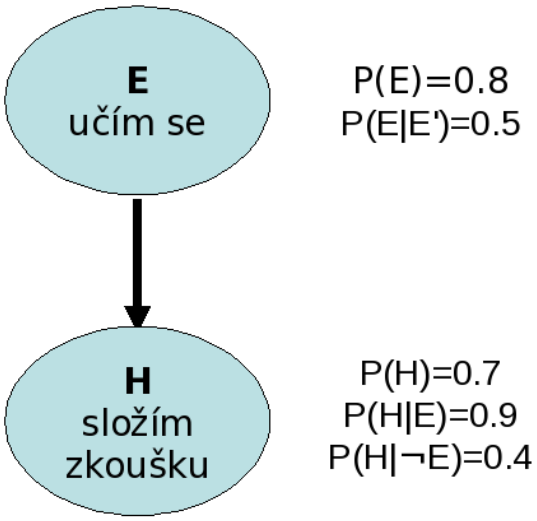
PROSPECTOR

- jeden z prvních úspěšně komerčně úspěšných **diagnostických** expertních systémů,
  - tvrzení dělí na (částečně) pozorovatelné předpoklady a diagnózy (závěry),
  - uvažování vždy ve směru od pozorování k závěrům,
  - if E then H with L (váha) else H with $\hat{L}$,
- geologie – vyhledávání nalezišť molybdenu, drahý hloubkový vrt?
  - 1978, nalezen zdroj v hodnotě 100 mil$
- **pseudopravděpodobnostní** model,
  - nekonzistentní systém měr – přezadaná přímka viz níže,
  - předpokládá stat. nezávislost předpokladů (rodičovských uzlů),
- řeší tři základní problémy (podobně jako ostatní systémy s neurčitostí)
  1. jak vyjádřit neurčitost elementárního tvrzení (předpokladu, E),
  2. jak aplikovat pravidlo pokud předpoklad není jistý,
  3. jak určit míru důvěry závěru (hypotézy, H), pokud je více pravidel se stejným závěrem.

PROSPECTOR

- ad 1) neurčitost předpokladů hodnotí expert pravděpodobností (subjektivně)
- ad 2) vychází z 5 veličin a grafického modelu uvedeného níže
  - $P(E)$, $P(H)$, $P(H|E)$, $P(H|\neg E)$ – popis vztahu předpoklad – závěr,
  - $P(E|E')$ – okamžité ohodnocení předpokladu,





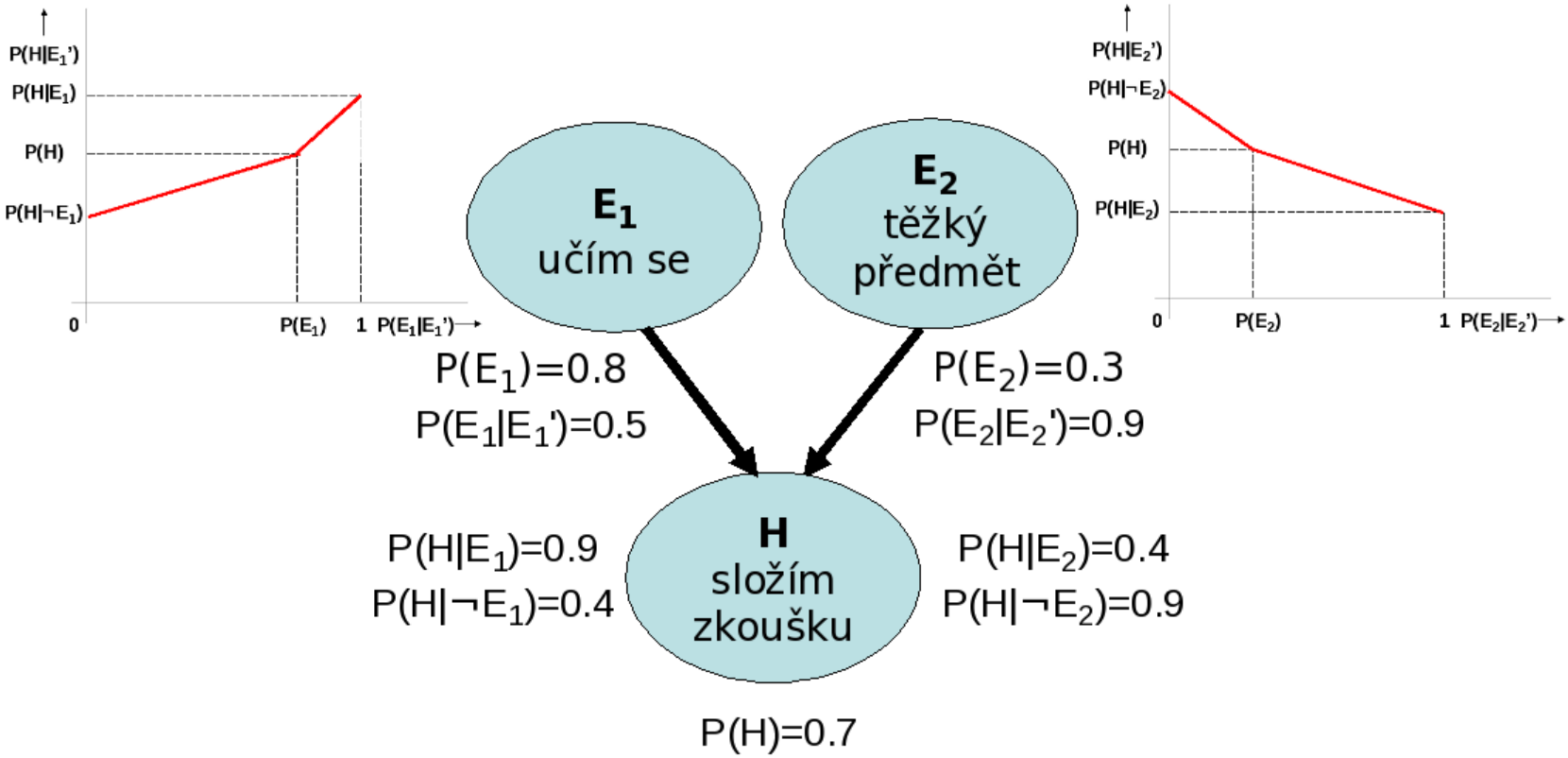
$$P(H|E') = P(H|\neg E) + (P(H) - P(H|\neg E))\frac{P(E|E')}{P(E)} = 0.4 + \frac{0.3 \times 0.5}{0.8} = 0.59$$

PROSPECTOR – skládání vlivu pozorování

- ad 3) pro kombinaci pozorování zavádí poměrnou pstní šanci (odds)
  - apriorní šance: $O(H) = \frac{P(H)}{P(\neg H)} = \frac{P(H)}{1-P(H)}$,
    - šance, že stihnu vlak je 1 ku 5 – jednou ano, pětkrát ne, pst $\frac{1}{6}$
  - aposteriorní šance: $O(H|E') = \frac{P(H|E')}{1-P(H|E')}$,
    - identická interpretace i pro podmíněné psti,
  - efektivní váha pravidla: $L' = \frac{O(H|E')}{O(H)}$,
    - $L' > 1$ ... předpoklad E' zvyšuje šanci splnění hypotézy,
- skládání pravidel je založeno na předpokladu nezávislosti
  - $O(H|E'_1, E'_2, \ldots, E'_n) = L'_1 \times L'_2 \times \cdots \times L'_n \times O(H)$
    - analogie bayesovského vztahu pro šance,
    - nová informace = informace z pozorování $\times$ původní informace,
  - zpětný převod: $P(H|E'_1, E'_2, \ldots, E'_n) = \frac{O(H|E'_1,E'_2,\ldots,E'_n)}{1+O(H|E'_1,E'_2,\ldots,E'_n)}$.

PROSPECTOR – skládání vlivu pozorování

- Co se stane pokud je předmět spíš těžký a učím se málo?
  - obě pozorování snižují apriorní pst úspěchu u zkoušky,
  - pst úspěchu poklesne z 0.7 na 0.33!



PROSPECTOR – skládání vlivu pozorování

- Apriorní šance: $O(H) = \frac{P(H)}{1-P(H)} = 2.\bar{3}$
- Efektivní váha pravidla o přípravě na zkoušku

  $P(H|E_1') = 0.59$ ... viz dříve

  $O(H|E_1') = \frac{P(H|E_1')}{1-P(H|E_1')} = 1.\bar{4}\bar{2}$

  $L_1' = \frac{O(H|E_1')}{O(H)} = 0.61$

- Efektivní váha pravidla o obtížnosti předmětu

  $P(H|E_2') = P(H) + (P(H) - P(H|\neg E_2))\frac{P(E_2|E_2')-P(E_2)}{1-P(E_2)} = 0.7 + \frac{-0.3\times 0.6}{0.7} = 0.44$

  $O(H|E_2') = \frac{P(H|E_2')}{1-P(H|E_2')} = 0.79$

  $L_2' = \frac{O(H|E_2')}{O(H)} = 0.34$

- Aposteriorní šance a pst

  $O(H|E_1', E_2') = L_1' \times L_2' \times O(H) = 0.48$

  $P(H|E_1', E_2') = \frac{O(H|E_1',E_2')}{1+O(H|E_1',E_2')} = 0.33$

MYCIN

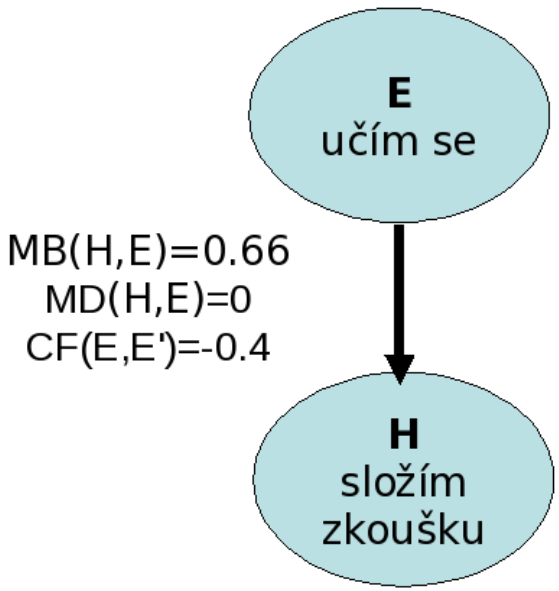
- 70. léta, Stanford University, programovací jazyk LISP,
- diagnostika infekčních onemocnění a doporučení léčby antibiotiky
  - 600 pravidel, úspěšnost vyšší než u lidských expertů,
  - přesto prakticky nenasazen – právní, etické a technické důvody,
- práce s neurčitostí založena na ad hoc **činiteli jistoty**,
- opět předpoklad statistické nezávislosti předpokladů,
- řeší tři základní problémy (podobně jako PROSPECTOR)
  1. jak vyjádřit neurčitost elementárního tvrzení (předpokladu, E),
  2. jak aplikovat pravidlo pokud předpoklad není jistý,
  3. jak určit míru důvěry závěru (hypotézy, H), pokud je více pravidel se stejným závěrem.

MYCIN

- ad 1) neurčitost předpokladů hodnotí expert **činitelem jistoty** CF (subjektivně)
  - $CF(E, E') \in \langle -1, 1 \rangle$, -1/1 $\sim$ předpoklad jistě nesplněn/splněn, 0 $\sim$ nevím
- ad 2) vychází ze 2 základních veličin
  - **míra důvěry** $MB(H, E)$ – kladná pokud splnění E vede ke vzrůstu důvěry v H,
    - $MB(H, E) \in \langle 0, 1 \rangle$, if $MD(H, E) > 0$ then $MB(H, E) = 0$
    - maximem je 1 (všechny objekty se splněným E mají splněn i závěr H),
  - **míra nedůvěry** $MD(H, E)$ – kladná pokud splnění E vede k poklesu důvěry v H,
    - $MD(H, E) \in \langle 0, 1 \rangle$, if $MB(H, E) > 0$ then $MD(H, E) = 0$
    - maximem je 1 (žádný objekt se splněným E nesplňuje závěr H)
  - důsledek – pokud E nemění důvěru experta v závěr H, pak $MB(H, E) = MD(H, E) = 0$
  - vztah mezi CF, MB a MD
    - MYCIN: $CF(H, E) = MB(H, E) - MD(H, E)$,
    - EMYCIN: $CF(H, E) = \frac{MB(H,E) - MD(H,E)}{1 - min(MB(H,E), MD(H,E))}$.

MYCIN – příklad

- Jak odvodit činitel jistoty závěru?
  - $MB(H, E') = MB(H, E) \times max(0, CF(E, E'))$,
  - $MD(H, E') = MD(H, E) \times max(0, CF(E, E'))$,
  - $CF(H, E') = MB(H, E') - MD(H, E')$.
- Příklad s přípravou na zkoušku
  - učení zvyšuje šanci u zkoušky,
    - $MB(H, E) > 0$ a $MD(H, E) = 0$,
  - včera jsem se učil méně než je obvyklé
    - $CF(E, E') < 0$,
  - šance na úspěch pozorováním nezměněna
    - $MB(H, E') = 0$ i $MD(H, E') = 0$
    - proto i $CF(H, E') = 0$.

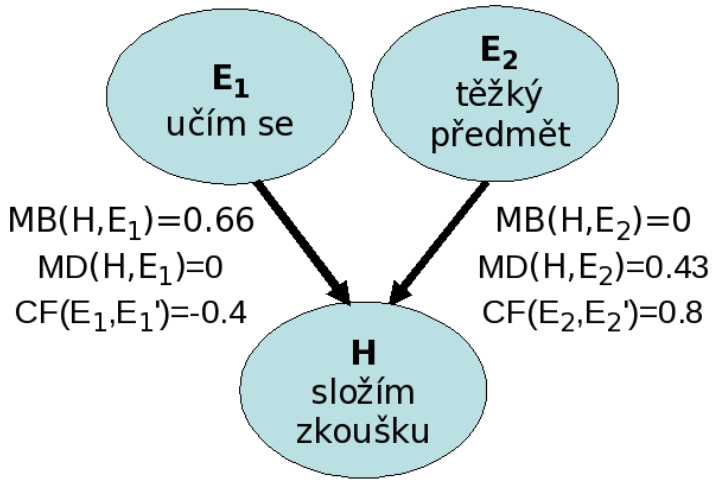


MYCIN – skládání vlivu pozorování

- ad 3) skládá pozorování (opakovanou) aplikací vztahů
  - $MB(H, E_1 \& E_2) = MB(H, E_1) + MB(H, E_2) - MB(H, E_1) \times MB(H, E_2)$,
  - $MD(H, E_1 \& E_2)$ analogicky,
- Stejný příklad: předmět je spíš těžký a učím se málo
  - první pozorování neovlivní CF (viz předchozí slajd),
  - druhé pozorování CF sníží
    - $CF(H, E_2') = -0.34$,
    - $MD(H, E_2') = MD(H, E_2) \times max(0, CF(E_2, E_2')) = 0.34$,
  - složení obou vlivů je triviální
    - $MB(H, E_1' \& E_2') = 0$,

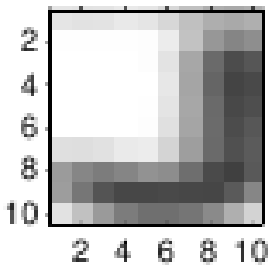
$$MD(H, E_1' \& E_2') = MD(H, E_1') + MD(H, E_2') - MD(H, E_1') \times MD(H, E_2') = \\ = MD(H, E_2') = 0.34$$

    - $CF(H, E_1' \& E_2') = -0.34$,



Srovnání s bayesovským usuzováním

- stejný cíl
  - změna psti platnosti hypotézy na základě pozorování
- založeno na Bayesově vztahu
  - $P(H|E) = \frac{P(E|H)P(H)}{P(E)}$, resp. $\frac{P(H|E)}{P(\neg H|E)} = \frac{P(H)}{P(\neg H)} \times \frac{P(E|H)}{P(E|\neg H)}$
- v kanonickém tvaru jednoduchý
  - předpokládá pozorování typu ano/ne,
  - k výpočetní explozi dochází, pokud je pozorování více,
  - často také pracuje s předpokladem nezávislosti,
- příklad z předmětu KUI
  - rozpoznání 20 různých znaků (A,B,...) $\longrightarrow$ 20 různých hypotéz,
  - na základě mřížky 10x10 bodů $\longrightarrow$ 100 pozorovatelných příznaků,
  - intenzitu každého bodu hodnotím na škále 0-255 $\longrightarrow$ analogie nejistoty,



  - potřebuji stanovit 20×256^{100} pstí typu $P(E_1 = e_1, E_2 = e_2, \ldots, E_{100} = e_{100}|H_A)$,
  - při nezávislosti bodů počet klesá na 20×256, stačí $P(E_1 = e_1|H_A)$, $P(E_2 = e_2|H_A)$, ...
  - další redukce lze dosáhnout slučováním intenzit apod.

Srovnání s bayesovským usuzováním

- PROSPECTOR
  - linearizuje pstní vztah mezi závěrem a předpokladem,
    - u příkladu se znaky pouze podmíněné psti pro nulovou a maximální intenzitu pixelu,
  - tím jej zjednodušuje a zpřehledňuje (pouze mezní hodnoty a apriorní psti),
  - linearizace je technicky nekonzistentní – rozklad na 2 části,
- MYCIN
  - je izomorfní, $MB(H,E)$ a $MD(H,E)$ lze vyjádřit jako fci $P(H|E)$ a $P(H)$,
  - $MB(H,E) = \frac{P(H|E)-P(H)}{1-P(H)}$, $MD(H,E) = \frac{P(H)-P(H|E)}{P(H)}$.
- je těžké nalézt výrazně odlišný přístup
  - to svědčí o jeho správnosti.

Shrnutí – modely neurčitosti ve znalostních systémech

- **výhody** určují praktickou použitelnost
  - stálý shell systému + problémově orientovaná znalostní báze,
  - výpočetně jednoduché, dostatečná výkonnost pro řadu úloh,
  - srozumitelnost pro lidi,
  - pružné hodnocení neurčitosti pozorování na kvantitativní škále,
- **nevýhody** jsou důsledkem zjednodušení
  - výhradně heuristické znalosti od expertů
    - výsledky přímo závisí na znalostech experta,
    - nijak neřeší problém učení inferenční sítě a pstí z dat,
    - v komplexních doménách je znalosti nemožné získat,
  - výpočetní jednoduchost vychází z ústupků
    - (podmíněná) nezávislost předpokladů,
    - struktura sítě předurčuje i směr usuzování,
    - způsob práce s neurčitostí hraje malou roli,
- spíše historie . . .
  - . . . v tomoto kurzu pro dokreslení principiálních přístupů k neurčitosti.

Doporučené doplňky – zdroje přednášky

:: Četba

- Russel, Norvig: **AI: A Modern Approach**, Uncertain Knowledge and Reasoning (Part V)
  - zejména neurčitost (kap. 14) a pravděpodobnostní usuzování (kap. 15)
- Mařík a kol. **Umělá inteligence 2**
  - kapitola Expertní systémy:
    diagnostické systémy s neurčitostí - MYCIN, PROSPECTOR, algebraická teorie,
  - kapitola Neurčitost v expertních systémech:
    typy neurčitosti, fuzzy míry, Dempster-Shaferova teorie, Bayesovské sítě,
- Mařík a kol. **Umělá inteligence 4**
  - kapitola Zobecněná teorie informace: teorie, formalizace a měření neurčitosti,
- Buchanan, Shortliffe: **Rule-Based Expert Systems**
  - podtitul: The MYCIN Experiments of the Stanford Heuristic Programming Project,
  - e-book, dostupná online.