



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Název projektu	Rozvoj vzdělávání na Slezské univerzitě v Opavě
Registrační číslo projektu	CZ.02.2.69/0.0./0.0/16_015/0002400

Expertní systémy

Neurčitost – Faktory jistoty

Jan Górecki



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Neurčitost (opakování)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Neurčitost je charakteristickým rysem složitých systémů. Vlastní povaha reality způsobuje, že poznatky, které z ní získáváme, jsou neurčité či vágní.
-

Příčiny neurčitosti (opakování)



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- problémy s daty; např.:
 - chybějící nebo nedostupná data
 - nespolehlivá data (např. z důvodu chyb měření)
 - nepřesná nebo nekonzistentní reprezentace dat
-

Příčiny neurčitosti (opakování)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- nejisté znalosti; např.:
 - znalost nemusí být platná ve všech případech
 - znalost může obsahovat vágní pojmy.
-

Vyjádření neurčitosti (opakování)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Neurčitost bývá v ES vyjadřována obvykle numerickými parametry, které se v různých systémech nazývají různě, např. *váhy*, *míry*, *stupně důvěry*, *faktory jistoty*.
-

Vyjádření neurčitosti (opakování)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Tyto numerické parametry se přiřazují jednotlivým tvrzením nebo pravidlům. Často nabývají hodnot z intervalu $\langle 0, 1 \rangle$ nebo $\langle -1, 1 \rangle$.
-

Vyjádření neurčitosti (opakování)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Většinou se neurčitost vyjadřuje pomocí jediného čísla. Postupně se však začínají prosazovat přístupy, v nichž je neurčitost vyjadřována dvojicí čísel (tato dvojice může být např. interpretována jako interval hodnot).
-

Problémy při zpracování neurčitosti



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Jak kombinovat neurčitá data v předpokladu pravidla?
 - Jak kombinovat neurčitost předpokladu pravidla a neurčitost pravidla jako celku?
-

Přístup založený na faktorech jistoty



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Faktory jistoty (*certainty factors*) byly poprvé použity v systému MYCIN. Cílem bylo eliminovat některé slabiny čistě pravděpodobnostního přístupu.
-

Přístup založený na faktorech jistoty



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Znalosti jsou vyjádřeny opět ve tvaru pravidel $E \rightarrow H$, přičemž s každým pravidlem je spojen faktor jistoty CF . Tento faktor nabývá hodnot z intervalu $\langle -1, 1 \rangle$ a je určen pomocí měr důvěry a nedůvěry MB a MD :

$$CF = \frac{MB - MD}{1 - \min\{MB, MD\}}$$

Přístup založený na faktorech jistoty



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Faktor jistoty vyjadřuje stupeň důvěry v hypotézu H , jestliže předpoklad E je pravdivý (1 znamená absolutní důvěru, -1 absolutní nedůvěru).
-

Míra důvěry



- Míra důvěry (*measure of belief*):

$$MB(H, E) = \begin{cases} 1 & \text{pro } P(H) = 1 \\ \frac{\max\{P(H | E), P(H)\} - P(H)}{1 - P(H)} & \text{jinak} \end{cases}$$

- Míra důvěry nabývá hodnot z intervalu $\langle 0, 1 \rangle$. Vyjadřuje stupeň, ve kterém je důvěra v hypotézu H podporována pozorováním evidence E .
-

Míra nedůvěry



- Míra nedůvěry (*measure of disbelief*):

$$MD(H, E) = \begin{cases} 1 & \text{pro } P(H) = 0 \\ \frac{P(H) - \min\{P(H | E), P(H)\}}{P(H)} & \text{jinak} \end{cases}$$

- Míra nedůvěry nabývá hodnot z intervalu $\langle 0, 1 \rangle$. Vyjadřuje stupeň, ve kterém je nedůvěra v hypotézu H podporována pozorováním evidence E .
-

Vliv neurčitosti předpokladů



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Předpoklad E v pravidle $E \rightarrow H$ nemusí být znám s absolutní jistotou. Může být odvozen z jiného pravidla nebo zadán uživatelem s nějakým faktorem jistoty $CF(E)$. Pak se výsledný faktor jistoty vypočte takto:

$$CF_{new}(H, E) = CF_{old}(H, E) \cdot CF(E)$$

Vliv neurčitosti předpokladů



- Jestliže přitom předpoklad E obsahuje konjunkci nebo disjunkci dílčích podmínek, pak se při výpočtu $CF(E)$ použijí následující vzorce:

$$CF(E_1 \wedge E_2) = \min\{CF(E_1), CF(E_2)\}$$

$$CF(E_1 \vee E_2) = \max\{CF(E_1), CF(E_2)\}$$

Kombinace více pravidel



Mějme pravidla $E_1 \rightarrow H, E_2 \rightarrow H, \dots, E_n \rightarrow H$. Označme

$CF_n = CF(H, E_1, \dots, E_n)$. Výpočet CF_n se provede podle následujícího

vzorce:

$$CF_n = \begin{cases} CF_{n-1} + CF(H, E_n) \cdot (1 - CF_{n-1}) & \text{pro } CF_{n-1} > 0 \text{ a } CF(H, E_n) > 0 \\ CF_{n-1} + CF(H, E_n) \cdot (1 + CF_{n-1}) & \text{pro } CF_{n-1} < 0 \text{ a } CF(H, E_n) < 0 \\ \frac{CF_{n-1} + CF(H, E_n)}{1 - \min\{|CF_{n-1}|, |CF(H, E_n)|\}} & \text{jinak} \end{cases}$$

Výhody a nevýhody faktorů jistoty



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

- Výhody:
 - jednoduchý a účinný výpočetní model
 - shromáždění potřebných dat podstatně snazší než v jiných metodách
 - snazší implementace vysvětlovacího mechanismu

 - Nevýhody:
 - chybí pevné teoretické základy
 - implicitní předpoklad nezávislosti evidencí E_i
-

Příklad (ES typu MYCIN)

příklad skutečné
diagnostiky spalovacího
čtyřtakového motoru

IF $\langle E_1 : \text{Klepou ventily po přidání plynu} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle MB(H, E_1) = 0,8; MD(H, E_1) = 0 \rangle$

IF $\langle E_2 : \text{Motor střílí do výfuku} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle MB(H, E_2) = 0,92; MD(H, E_2) = 0 \rangle$

IF $\langle E_3 : \text{Malá spotřeba} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle MB(H, E_3) = 0; MD(H, E_3) = 0,9 \rangle$

IF $\langle E_4 : \text{Malá akcelerace} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle MB(H, E_4) = 0,86; MD(H, E_4) = 0 \rangle$



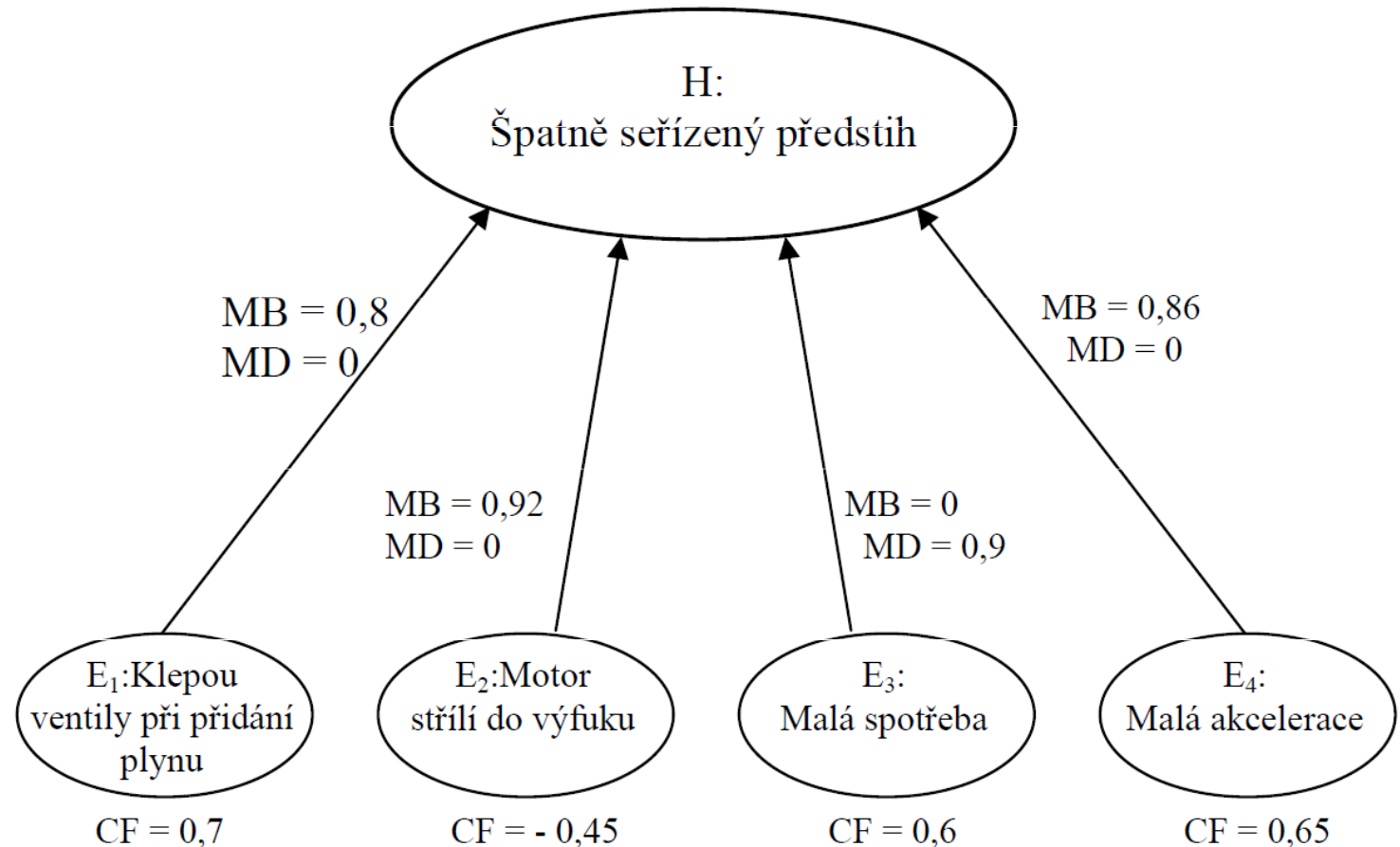
Příklad (ES typu MYCIN)

IF $\langle E_1 : \text{Klepou ventily po přidání plynu} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle MB(H, E_1) = 0,8; MD(H, E_1) = 0 \rangle$

IF $\langle E_2 : \text{Motor střílí do výfuku} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle MB(H, E_2) = 0,92; MD(H, E_2) = 0 \rangle$

IF $\langle E_3 : \text{Malá spotřeba} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle MB(H, E_3) = 0; MD(H, E_3) = 0,9 \rangle$

IF $\langle E_4 : \text{Malá akcelerace} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle MB(H, E_4) = 0,86; MD(H, E_4) = 0 \rangle$



Příklad (ES typu MYCIN)



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Předpokládejme nyní, že uživatel posoudí projevy diagnostikovaného motoru a ohodnotí platnost evidencí E_1 až E_4 postupně takto:

$$CF(E_1, E') = 0,7 ,$$

$$CF(E_2, E') = -0,45 ,$$

$$CF(E_3, E') = 0,6 ,$$

$$CF(E_4, E') = 0,65 ,$$

Příklad (ES typu MYCIN)



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Pro ohodnocení první evidence platí

$$MB(H, E'_1) = MB(H, E_1) \cdot \max\{0, CF(E_1, E')\} = 0,8 \cdot 0,7 = 0,56$$

$$MD(H, E'_1) = 0$$

$$CF(H, E'_1) = 0,56$$

Příklad (ES typu MYCIN)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Po určení $CF(E_2, E') = -0,45$ ➡➡

$$MB(H, E'_2) = MD(H, E'_2) = CF(H, E'_2) = 0$$

protože pravidlo nepřispívá k ohodnocení H a proto

$$CF(H, E'_1 \& E'_2) = CF(H, E'_1) = 0,56$$

Příklad (ES typu MYCIN)



Dále uživatel odhadl $CF(E_3, E') = 0,6$, potom

$$MB(H, E'_3) = 0$$

$$MD(H, E'_3) = MD(H, E_3) \cdot \max\{0, CF(E_3, E')\} = 0,9 \cdot 0,6 = 0,54$$

$$CF(H, E'_3) = -0,54$$

$$\begin{aligned} CF(H, E'_1 \& E'_2 \& E'_3) &= \frac{CF(H, E'_1 \& E'_2) \uparrow CF(H, E'_3)}{1 - \min\{\text{abs}(CF(H, E'_1 \& E'_2)), \text{abs}(CF(H, E'_3))\}} = \\ &= \frac{0,56 - 0,54}{1 - 0,54} = 0,0434 \end{aligned}$$

Příklad (ES typu MYCIN)



V posledním kroku pak vypočteme

$$MB(H, E'_4) = MB(H, E_4) \cdot \max\{0, CF(E_4, E')\} = 0,86 \cdot 0,65 = 0,559$$

$$MD(H, E'_4) = 0$$

$$CF(H, E'_4) = 0,559$$

$$\begin{aligned} CF(H, E'_1 \& E'_2 \& E'_3 \& E'_4) &= CF(H, E'_1 \& E'_2 \& E'_3) + CF(H, E'_4) - \\ &\quad - CF(H, E'_1 \& E'_2 \& E'_3) \cdot CF(H, E'_4) = \\ &= 0,578 \end{aligned}$$

Příklad (ES typu MYCIN)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Povšimneme si, že pozitivní odpovědi uživatele na dotazy

„Klepou ventily při přidání plynu ?“ a *„Má motor malou akceleraci ?“*

podporovaly podezření, že je

„Špatně seřízený předstih“.

Naopak pozitivní odpověď na otázku

„Má motor malou spotřebu ?“

pak podezření na tuto diagnózu zeslabuje (vzhledem k nenulové MD v
a negativní odpověď na dotaz

„Střílí motor do výfuku ?“

v našem modelu vůbec ohodnocení hypotézy H neovlivní.

Příklad (ES typu PROSPECTOR)



IF $\langle E_1 : \text{Klepou ventily při přidání plynu} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle P(H | E_1) = 0,85; P(H | \text{not } E_1) = 0,05 \rangle$

IF $\langle E_2 : \text{Motor střílí do výfuku} \rangle$
THEN $\langle H : \text{Špatně seřízený předstih} \rangle$
WITH $\langle P(H | E_2) = 0,95; P(H | \text{not } E_2) = 0,3 \rangle$

Příklad (ES typu PROSPECTOR)



IF $\langle E_3 : \text{Malá spotřeba} \rangle$

THEN $\langle H : \text{Špatně seřízený předstih} \rangle$

WITH $\langle P(H | E_3) = 0,001; P(H | \text{not } E_3) = 0,7 \rangle$

IF $\langle E_4 : \text{Malá akcelerace} \rangle$

THEN $\langle H : \text{Špatně seřízený předstih} \rangle$

WITH $\langle P(H | E_4) = 0,9; P(H | \text{not } E_4) = 0,1 \rangle$

Příklad (ES typu PROSPECTOR)



Předpokládejme, že jsme od experta získali následující apriorní pravděpodobnosti:

$$P(E_1) = 0,1; P(E_2) = 0,05; P(E_3) = 0,6; P(E_4) = 0,4 \text{ a } P(H) = 0,35.$$

Uživatel postupně zadává odpovědi:

a) $P(E | E'_1) = 0,8$, pak

$$P(H | E'_1) = P(H) + \frac{P(H | E_1) - P(H)}{1 - P(E_1)} \cdot (P(E_1 | E'_1) - P(E_1)) = 0,739$$

$$L'_1 = \frac{O(H | E'_1)}{O(H)} = \frac{\frac{P(H | E'_1)}{1 - P(H | E'_1)}}{\frac{P(H)}{1 - P(H)}} = 5,258$$

přičemž

$$O(H) = 0,5385; O(H | E'_1) = 2,831$$

Příklad (ES typu PROSPECTOR)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

b) $P(E | E'_2) = 0,01$, pak

$$P(H | E'_2) = P(H | \text{not } E_2) + \frac{P(H) - P(H | \text{not } E_2)}{P(E_2)} \cdot P(E_2 | E'_2) = 0,301$$

$$L'_2 = \frac{O(H | E'_2)}{O(H)} = 0,8$$

a tedy

$$O(H | E'_1, E'_2) = L'_1 \cdot L'_2 \cdot O(H) = L'_2 \cdot O(H | E'_1) = 0,8 \cdot 2,831 = 2,265$$

Odtud

$$P(H | E'_1, E'_2) = 0,694$$

Příklad (ES typu PROSPECTOR)



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

c) $P(E | E'_3) = 0,071$, pak analogicky jako v předchozích případech vypočítáme

$$P(H | E'_3) = 0,071$$

$$L'_3 = 0,142$$

$$\begin{aligned} O(H | E'_1, E'_2, E'_3) &= L'_1, L'_2, L'_3 \cdot O(H) = L'_3 \cdot O(H | E'_1, E'_2) = \\ &= 0,142 \cdot 2,265 = 0,322 \end{aligned}$$

$$P(H | E'_1, E'_2, E'_3) = 0,244$$

Příklad (ES typu PROSPECTOR)



d) po odpovědi na poslední otázku uvede $P(E | E'_4) = 0,8$, pak

$$P(H | E'_4) = 0,717$$

$$L'_4 = 4,705$$

$$O(H | E'_1, E'_2, E'_3, E'_4) = L'_4 \cdot O(H | E'_1, E'_2, E'_3) = 4,705 \cdot 0,224 = 1,148$$

$$P(H | E'_1, E'_2, E'_3, E'_4) = 0,534$$

Děkuji za pozornost

Některé snímky převzaty od:

RNDr. Jiří Dvořák, CSc. dvorak@fme.vutbr.cz

prof. Dr. Ing. Miroslav Pokorný miroslav.pokorny@vsb.cz