

Případové usuzování v expertním systému NEST

Vladimír Laš , Petr Berka

Katedra informačního a znalostního inženýrství, FIS, VŠE – Praha
Praha 3, Nám. W. Churchilla 4
{lasv,berka}@vse.cz

Abstrakt. Článek popisuje základní způsob případového usuzování, který byl implementován do expertního systému NEST, původně pouze pravidlového. Odvozování probíhá nad skladem předchozích případů, který (na rozdíl od pravidlové báze znalostí) nemusí vytvářet expert. V systému jsou implementovány dva způsoby případového odvozování – kompozicionální a logický – lišící se způsobem využívání případů ve skladu. Tyto způsoby jsou inspirovány kompozicionálním a logickým způsobem práce s neurčitostí v síti pravidel původního systému NEST. V závěru jsou nastíněny další možné rozšíření systému – především hlubší integrace pravidlového a případového usuzování.

Klíčová slova: případové usuzování, expertní systém

1 Úvod

Tento text popisuje, jakým způsobem funguje případové usuzování, které bylo přidáno do expertního systému NEST.

Odvozování (kompozicionální respektive logické) v tomto expertním systému bylo původně založeno pouze na pravidlech (klasická IF THEN pravidla). Důvodů pro přidání případového usuzování do systému je několik. Mezi hlavní patří problémy s tvorbou bází znalostí. Pro pravidlové usuzování je nutné mít předem vytvořenou kvalitní bázi znalostí, podle níž systém odvozuje závěry. Pro případové usuzování nám stačí mít sklad případů s dostatečným počtem dříve vyřešených případů (o kterých máme alespoň určitou jistotu, že byly řešeny správně). Tento sklad případů lze sehnat jednodušším způsobem než kvalitní bázi znalostí. Dalším důvodem je např. možnost odvodit výsledky různými způsoby a tím ověřit jejich správnost.

2 Systém NEST

V tomto textu nebudeme popisovat NEST, o tomto systému se lze více dozvědět v [1] či [2]. Zde bychom zmínili pouze dvě základní věci, které přímo souvisí s případovým usuzováním – reprezentace znalostí pomocí výroků a pravidel a odvozování a způsob práce s neurčitostí .

2.1 Výroky a pravidla

Systém NEST nabízí pro reprezentování znalostí výroky (odvozené od binárních, nominálních a numerických atributů), pravidla (kompozicionální s neurčitostí i nekompozicionální bez neurčitosti), kontexty a integritní omezení. IF-THEN pravidla představují původní základní prostředek pro reprezentování znalostí v systému NEST. Předpoklad pravidla je tvořen disjunktivní normální formou (disjunkce konjunkcí) literálů (výroků nebo jejich negací), závěr pravidla je tvořen seznamem (konjunkcí) literálů. Systém NEST nabízí tři typy pravidel:

- *kompozicionální pravidla* jsou pravidla, kde každý závěr pravidla je doplněn váhou pravidla, která zhruba řečeno vyjadřuje do jaké míry tento závěr platí, pokud předpoklad platí s jistotou. V průběhu konzultace se skládají dílčí příspěvky pravidel, která vedou ke stejnému závěru – proto označení kompozicionální,
- *apriorní pravidla* jsou kompozicionální pravidla bez předpokladu; tato pravidla umožňují přiřadit výroky implicitní váhu,
- *logická pravidla* jsou nekompozicionální pravidla bez vah. Je-li předpoklad takového pravidla s jistotou splněn, váha závěru je známa rovněž s jistotou a závěr se považuje za plně vyhodnocený – v ostatních případech se pravidlo neaktivuje. Pouze logická pravidla mohou dát absolutní jistotu $+1$.

2.2 Odvozování a práce s neurčitostí

Systém NEST umožňuje pracovat s váhami, které vyjadřují neurčitost v bázi znalostí i neurčité odpovědi během konzultace. Práce s neurčitostí vychází z algebraické teorie P. Hájka [3]. V systému jsou tedy definovány kombinační funkce NEG (pro výpočet váhy negace výroku), CONJ (pro výpočet váhy konjunkce výroků), DISJ (pro výpočet váhy disjunkce výroků), CTR (pro výpočet příspěvku pravidla k váze závěru) a GLOB (pro složení příspěvků více pravidel se stejným závěrem). Váha není vyjádřena jedním číslem, ale reprezentuje ji interval vah. Interval chápeme tak, že výrok může získat libovolnou váhu z daného intervalu. Interval tedy simuluje opakované konzultace pro všechny možné hodnoty z intervalu. Vzhledem k monotonii kombinačních funkcí (funkcí používaných při odvozování) stačí pracovat pouze s mezemi intervalů.

Použité typy pravidel nabízejí dva možné způsoby odvozování:

- *kompozicionální odvozování* v rámci kompozicionálních a apriorních pravidel. V průběhu konzultace se skládají dílčí příspěvky všech pravidel vedoucích ke stejnému závěru. Některé pravidlo je pak aktivováno a dá nenulový příspěvek k váze závěru, pokud (1) kontext pravidla (je-li zadán) je odvozen s kladnou vahou a rovněž (2) předpoklad pravidla je odvozen s kladnou vahou – kontext (je-li zadán) přitom modifikuje váhu pravidla. Jde tedy o způsob odvozování známý ze systému SAK (a analogický systémům Equant, FEL-EXPERT, MYCIN, PROSPECTOR a pod.),
- *nekompozicionální odvozování* v rámci logických pravidel. Jak již bylo uvedeno v předcházejícím odstavci, logická pravidla odvozují závěr

s jistotou – s využitím klasického dedukčního pravidla modus ponens. V průběhu konzultace pak jedno aktivované pravidlo plně vyhodnotí daný závěr. “Skládání” příspěvků pravidel tedy má charakter disjunkce, tak jak je tomu například v komerčních expertních systémech Nexpert Object nebo Kappa PC.

Z tohoto způsobu rozdělení odvozování na kompozicionální a logické vychází i způsob práce případového usuzování, jak je uvedeno dále.

3 Jak funguje případové usuzování

Případové usuzování (Case-Based Reasoning, CBR) je založeno na principu analogie: v neznámé situaci použij to řešení, které se osvědčilo v situaci podobné. Znalosti jsou tedy reprezentovány v podobě databáze již vyřešených problémů (případů) a usuzování má podobu hledání případu, který bude nejpodobnější nastalé rozhodovací situaci [4].

Případy vyjadřující znalosti v systémech CBR lze reprezentovat v zásadě dvěma způsoby:

1. jako nestruturované vektory hodnot atributů,
2. jako strukturované objekty, nesoucí informace o situacích, kdy je použití případu relevantní i o akcích, které se mají provést.

Použitý způsob reprezentace pak předurčuje způsob ukládání (indexování) případů i měření podobnosti. V prvním případě vystačíme s některou z běžných metrik. Pro měření vzdálenosti mezi numerickými hodnotami se obvykle používá eukleidovská metrika

$$d_E(x_1, x_2) = \sqrt{\sum_{j=1}^m \delta_E(x_{1j}, x_{2j})}, \quad \text{kde } \delta_E(x_{1j}, x_{2j}) = (x_{1j} - x_{2j})^2 \quad (1)$$

pro měření vzdálenosti mezi kategoriálními hodnotami se používá tzv. překrytí (overlap)

$$d_O(x_1, x_2) = \sum_{j=1}^m \delta_O(x_{1j}, x_{2j}),$$

$$\text{kde: } \begin{aligned} \delta_O(x_{1j}, x_{2j}) &= 0 \text{ pro } x_{1j} = x_{2j} \\ \delta_O(x_{1j}, x_{2j}) &= 1 \text{ pro } x_{1j} \neq x_{2j} \end{aligned} \quad (2)$$

kde x_1, x_2 jsou dva vektory o m složkách.

V systému NEST jsme zvolili první způsob reprezentace (a tedy i odvozování). Případy jsou chápány jako údaje z provedených konzultací: obsahují tedy informace o odpovědích na dotazy i o odvozených cílech. Odvozování vah závěrů nové konzultace pak probíhá na základě dřívějších případů. K tomu je zapotřebí, abychom měli sklad případů s dostatečným počtem dříve vyřešených případů. Nová konzultace (nový případ) je pak porovnávána s těmito staršími případy, je určena podobnost

k jednotlivým případům a na základě této podobnosti je spočtena váha cílů nové konzultace.

Podobně jako v případě použití pravidel, i v případě použití případů nabízí NEST dva přístupy: kompozicionální odvozování a logické odvozování. Oba způsoby definují 3 základní funkce:

- Podobnost dvou výroků – výsledkem je číslo z intervalu $[-1;1]$, které určuje, jak „podobné“ jsou dva stejné výroky z různých konzultací (při práci s intervaly musíme počítat s intervalem $[w_1, w_2]$, kde w_i je prvkem $[-1;1]$).
- Podobnost dvou konzultací – výsledkem je opět číslo z intervalu $[-1;1]$, které určuje, jak „podobné“ jsou dvě konzultace. Funkce pro výpočet využívá vzdálenost jednotlivých výroků.
- Výpočet váhy závěrů - je funkce, která na základě vzdálenosti 2 konzultací (vzdálenost nové konzultace k jednotlivým starým případům) vypočítá váhy závěrů nové konzultace.

Základní rozdíl mezi kompozicionálním a logickým způsobem odvozování spočívá v tom, že zatímco v kompozicionálním způsobu jsou brány v úvahu váhy všech případů ve skladu a váha závěru je odvozována skládáním těchto vah, tak v logickém způsobu odvozování je vybrán nejpodobnější případ a váha závěru je určena na základě tohoto případu.

Cíle (cílové výroky) nové konzultace mohou být určeny dvěma způsoby. Pokud v bázi znalostí existují pravidla, mohou být cíle určeny na základě těchto pravidel stejným způsobem jako v případě pravidlového usuzování. Druhou možností je přímo ve skladu případů označit atributy, které mají být brány jako cílové.

Jednotlivé atributy mohou mít ve skladu případů určeny svoji váhu. Tato váha je pak zohledňována při výpočtu vzdálenosti 2 konzultací (atributy s vyšší váhou se na odvozování podílejí více než atributy s nižší váhou). Pokud atribut nemá váhu zadanou, bere se automaticky maximální váha (tedy 1).

4 Kompozicionální odvozování

4.1 Výpočet podobnosti dvou výroků

Mějme výrok v , který má v konzultaci k_a váhu w_a a v konzultaci k_b váhu w_b . Podobnost výroku pak definujeme jako:

$$pv_k(k_a, k_b) = 1 - |w_a - w_b| \quad (3)$$

$$pv_k \in [-1;1]$$

$pv_k = -1$ znamená, že výroky jsou zcela odlišné, $pv_k = 1$ znamená, že výroky jsou stejné.

Pro práci s intervaly (uvažujme intervaly $[w_{a1}, w_{a2}]$ a $[w_{b1}, w_{b2}]$): Jestliže $aw_1 > bw_2$ nebo $bw_1 > aw_2$ (intervaly se nepřekrývají) pak

$$pv_k(k_a, k_b) = [\min(1 - |w_{a1} - w_{b2}|, 1 - |w_{a2} - w_{b1}|), \max(1 - |w_{a1} - w_{b2}|, 1 - |w_{a2} - w_{b1}|)] \quad (4)$$

Jinak (intervaly se překrývají):

$$x_1 = |w_{a1} - w_{b1}|, x_2 = |w_{a1} - w_{b2}|, x_3 = |w_{a2} - w_{b1}|, x_4 = |w_{a2} - w_{b2}| \quad (5)$$

$$pv_k(k_a, k_b) = [1 - \max(x_1, x_2, x_3, x_4), 1]$$

4.2 Výpočet podobnosti dvou konzultací

Mějme konzultaci k_1 a konzultaci k_2 . Podobnost konzultací pak definujeme jako:

$$pk(k_1, k_2) = \frac{\sum_i pv(v_{k1i}, v_{k2i}) * \frac{wa}{pva}}{sva} \quad (6)$$

kde

- v_{kli} značí i-tý výrok (dotaz) v případě k_i
- wa je váha atributu nastavená v bázi znalostí,
- pva je počet výroků, které jsou odvozeny od rodičovského atributu příslušného výroku (toto dělení je zde proto, neboť každý atribut má být při výpočtu podobnosti zohledňován jen jednou; je-li tedy od atributu odvozeno více výroků, je tímto způsobem celkově započítán jen jednou)
- sva je součet vah atributů (dotazů)
- Σ probíhá přes všechny dotazy

$pk = -1$ znamená, že případy jsou zcela odlišné, $pk = 1$ znamená, že případy jsou stejné.

Pro práci s intervaly se podle vzorce zvlášť spočte hodnota pro horní mez a zvlášť pro dolní mez.

4.3 Výpočet váhy závěru nové konzultace

Mějme bázi znalostí s případy k_i a novou konzultaci (nový případ) k_n . Mějme cílový výrok c s vahou w_c

$$w_{cn} = \frac{\sum_i pk(k_n, k_i) \cdot w_{ci}}{pkon} \quad (7)$$

kde

- $pkon$ je počet konzultací
- Σ jde přes všechny případy v bázi znalostí.

Jedná se vlastně o vážený součet vah závěrů jednotlivých případů, kde jako váhy slouží podobnost jednotlivých případů od nové konzultace.

Pro práci s intervaly se podle vzorce zvlášť spočte hodnota pro horní mez a zvlášť pro dolní mez.

5 Logické odvozování

5.1 Výpočet podobnosti dvou výroků

Mějme výrok v , který má v konzultaci k_1 váhu w_1 a v konzultaci k_2 váhu w_2 . Dále mějme zadaný práh podobnosti výroků p . Podobnost výroků pak definujeme jako:

$$\begin{aligned}pv_i(k_1, k_2) &= 1 \text{ jestliže } pv_k(k_1, k_2) > p \\pv_i(k_1, k_2) &= -1 \text{ jinak}\end{aligned}\tag{8}$$

To znamená, že pokud vzdálenost výroků spočtena kompozicionálním přístupem přesáhne práh podobnosti výroků, pak je logická podobnost rovna 1, jinak se rovná -1. Pro práci s intervaly se spočte stejným způsobem pro minimální i maximální váhu.

5.2 Výpočet podobnosti dvou konzultací

Počítá se stejným způsobem jako u kompozicionálního přístupu.

5.3 Výpočet váhy závěru nové konzultace

Mějme bázi znalostí s případy k_i a novou konzultaci k_n . Mějme cílový výrok c s vahou w_c

$$w_{cn} = \frac{\sum_{i \in M} w_{ci}}{|M|} : i \in M \Leftrightarrow i = \max_i [pk(k_n, k_i)]\tag{9}$$

Nebo-li váha nové konzultace se rovná váze konzultace, která je nejpodobnější a v případě, že je nejpodobnějších konzultací více, vezme se průměr z jejich vah.

Pro práci s intervaly je případ s podobností $[a_1, a_2]$ podobnější nové konzultaci než případ s podobností $[b_1, b_2]$ jestliže:

$$(a_1 + a_2) - (b_1 + b_2) > 0\tag{10}$$

6 Implementace systému

Implementace pravidlového usuzování systému plynule navazuje na původní implementaci systému, jak je popsána v [1]. Zatím je vytvářena pouze stand-alone verze (klient-server verze bude dodělána až pod odladění všech problémů).

Sklad případů je nyní reprezentován pomocí souboru xml (obdobně jako báze znalostí s pravidly), přičemž ukáže-li se v budoucnu, že případové odvozování nad tímto souborem probíhá příliš pomalu, bude vytvořen vlastní formát pro skladování dat (případně bude použita relační databáze).

7 Další rozvoj systému

V prvním kroku rozvoje systému je potřeba především sehnat dostatečně obsáhlý sklad případů a vytvořit pravidlovou bázi znalostí z příslušné oblasti. Na těchto základech by mělo proběhnout důkladné testování výše popsaných algoritmů.

V dalším rozvoji by mělo dojít k užšímu propojení případového a pravidlového usuzování, kdy by systém měl automaticky (či poloautomaticky) vybírat mezi případovým a pravidlovým usuzováním.

V poslední části rozšíření bude snaha naučit systém upravovat pravidla v bázi znalostí na základě skladu případů.

Reference

1. Berka P., Laš V., Svátek V., Nový expertní systém NEST. Brno 25.02.2004 – 27.02.2004. In: Snášel V. (ed.). *Znalosti 2004*. Ostrava : VŠB TU Ostrava, 2004, s. 256–267. ISBN 80-248-0456-5
2. Laš V. <http://lisp.vse.cz/NEST>, Expertní systém NEST
3. Hájek P. Combining Functions for Certainty Factors in Consulting Systems. *Int. J. Man-Machine Studies* 22, 1985
4. Watson I., Marir F., Case-based reasoning: An review. *The Knowledge Engineering Review*, Vol. 9:4, 1994, 327-354.

Annotation:

This paper describes basic type of case-based reasoning how it was implemented in the expert system NEST (originally rule-based system). The reasoning proceeds over the cases storage, which doesn't have to be created by expert contrary to rule knowledge base. There are 2 types of case-based reasoning in the system – compositional and logical. These types differ in manner of using cases storage. Future developments of the system is outlined in the end of paper.