

Pokroky matematiky, fyziky a astronomie

Petr Jirků
Expertní systémy

Pokroky matematiky, fyziky a astronomie, Vol. 31 (1986), No. 3, 132--144

Persistent URL: <http://dml.cz/dmlcz/138391>

Terms of use:

© Jednota českých matematiků a fyziků, 1986

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

Postoj Strouhalův a jeho následovníků je do značné míry typický i pro Jednotu na konci minulého a na počátku tohoto století. Stála na pokrokových pozicích, které si i nadále uchovala.

Vývoj Jednoty samozřejmě nebyl hladký. Po období rozkvětu přišla i obtížnější léta. Vcelku však počet jejích členů neustále vzrůstal a tomu se přizpůsobovala i její struktura. V roce 1912 se mění její název na Jednotu českých matematiků a fyziků. Již v roce 1913 vzniká její pobočka v Brně a později se počet jejích poboček rozšiřuje. V roce 1919 získává Jednota vlastní tiskárnu, známý Prométheus, a buduje podnik schopný moderní a obtížné matematické a fyzikální sazby.

Velmi těžkou dobou pro Jednotu byla léta druhé světové války, která si právě v těchto dnech v duchu znovu revokujeme. Po nacistickém uzavření vysokých škol se Jednota všestranně snažila udržovat úroveň znalostí a informací ve fyzice a v matematice na žádoucí výši. Dělala to nejen přednáškami, ale i vydáváním sbírky Cesta k vědění. V roce 1941 byla však donucena přerušit vydávání Časopisu pro pěstování matematiky a fyziky; v r. 1943 musela uzavřít své knihkupectví. V roce 1944 je zatčen její předseda, vynikající fyzik, profesor Univerzity Karlovy F. Záviška, který na sklonku války v roce 1945 umírá na pochodu smrti z koncentračního tábora. S ním sdílí podobný osud i řada dalších členů Jednoty.

Jednota v současné době plní své tradiční úkoly. Napomáhá modernizaci a zdokonalování výuky matematiky a fyziky na základních, středních i vysokých školách, pečuje o realizaci nového výchovně vzdělávacího projektu a pracuje na perspektivních záměrech v této oblasti. Metodami vědeckých společností přispívá k rozvoji vědeckého života v matematice a fyzice. Sdružuje dnes spolu s JSMF téměř 7 000 členů, zaměřuje jejich aktivitu k tomu, aby přispěli k přenosu nových poznatků těchto oborů co nejrychleji do společenské praxe, a tím napomáhá rozvíjet naši socialistickou zemi.

Expertní systémy

Petr Jirků, Praha

1. Úvod

Expertní systémy jsou počítačové programy založené na znalostech a určené k mechanizování některých lidských kognitivních aktivit. Tyto programy mají uživateli umožnit dosáhnout vysoce kvalifikovaných znaleckých odpovědí na otázky z určité speciální oblasti. Obvykle jsou realizovány jako automatizované konzultační systémy. Problematika tvorby expertních systémů patří v posledních letech mezi bouřlivě se rozvíjející

odvětví umělé inteligence [1]. V tomto článku chceme seznámit čtenáře se současným stavem rozvoje výzkumu expertních systémů a současně poukázat na matematické problémy, které tvoří pozadí jejich výstavby. V tomto širším kontextu jde o *znalostní inženýrství*, tj. o zkoumání jednak struktury znalostí samých, jednak i procesů, jako je shromažďování, uchovávání a odvozování znalostí o vybrané speciální oblasti zájmu. Znalostní inženýrství se pak zaměřuje na to, jak psát počítačové programy, které mohou řešit problémy dosud tradičně považované za problémy vyžadující lidskou inteligenci.

Ve vědeckém výzkumu dlouho bylo (a zčásti možná ještě je) často nevyřčeným ideálem nalézání kvantitativních charakteristik jevů, pokud možno v podobě funkčních závislostí. R. Duda v jednom ze svých výkladů o expertních systémech [4] cituje v této souvislosti lorda Kelvina: „Můžete-li měřit to, o čem mluvíte, a můžete-li to vyjádřit číselně, něco o tom víte, avšak když to nemůžete měřit, tak vaše znalost je chudá a nedostatečná. Může to být pouze počátek poznání, ale stěží vědecký pokrok.“

To je ostře vyjádřený názor, který je bezpochyby silně dobově podmíněný. Z dnešního pohledu nám na něm vadí především to, že vylučuje z vědeckého poznávání i tak důležité postupy, jako řešení problémů, dokazování teorémů, plánování akcí, dedukci a další. Dnes můžeme dokonce říci, že většina poznávacích procesů se odehrává spíše v symbolické než v numerické rovině. A netýká se to pouze kognitivních postupů samotných, i v matematice jsou oblasti, v nichž nám jde spíše o kvalitativní než kvantitativní poznatky. A to nemáme na mysli pouze obory jako algebra, logika apod., v nichž zpracování symbolické informace je základem, ale např. i obory, v nichž není numerické vyjádření a priori vyloučeno. Příkladem může být třeba kvalitativní řešení diferencíálních rovnic. Jde spíše o kladení otázek určitého typu. A právě expertní systémy jsou založeny převážně na zpracování symbolické informace.

V současné době se odhaduje, že v běžném použití je více než jeden a půl tisíce expertních systémů, z toho více než 300 v oblasti biomedicínské. Jejich počet velmi rychle narůstá. To je mj. způsobeno tím, že vznikají problémově nezávislé (prázdné) expertní systémy, které jsou využívány v nejrozmanitějších aplikačních oblastech tak, že uživatelé je naplňují problémově orientovanými znalostmi z oboru jejich zájmu. Právě tato možnost, oddělit řídicí mechanismus expertního systému od vlastní báze znalostí, podnítila jejich překotné šíření.

Klasickým příkladem expertního systému je program MYCIN, navržený E. Shortlifem [15] pro určování diagnózy jistých infekčních onemocnění na základě bakteriálních vyšetření spolu se současným návrhem vhodné terapie odpovídajícím druhem antibiotik. MYCIN je realizován jako zpětný inferenční systém (viz dále), který od cílů (možných diagnóz) postupuje v kontextovém stromu znalostí (pravidel) vstříc klinickým nálezům a laboratorním výsledkům, aby našel argumenty podporující některý z možných závěrů, tj. aby určil nejpravděpodobnější druh onemocnění. Kontextovým stromem znalostí je specifikována řídicí strategie systému. Program může zpracovávat nejistou i neúplnou informaci. Nejistota je vyjadřována pomocí tzv. faktorů věření (faktům či znalostem), které jsou na nejnižší úrovni (na úrovni nálezů a laboratorních dat) přiřazovány heuristicky a pak propagovány sítí pravidel (často též heuristických) k určení hodnot cílů. MYCIN může navíc podávat vysvětlení svých rozhodnutí jednak pomocí kontextového stromu, jednak pomocí jednotlivých pravidel vedoucích k danému závěru. Na základě

systemu MYCIN byl vyvinut problémově nezávislý expertní systém EMYCIN (empty-MYCIN) [9], který si z MYCINu ponechal způsob vyjadřování a rozčlenění znalostí, způsob práce s nejistými znalostmi i základní řídicí strukturu, ale „zapomněl“ konkrétní znalosti, se kterými MYCIN pracoval. EMYCIN byl tedy připraven stejným způsobem zpracovávat i znalosti z jiných oborů. Byl již úspěšně využit i v oblastech mimo medicínu a stal se jakýmsi prototypem problémově nezávislých systémů.

Poněkud odlišnou kategorií expertních systémů charakterizuje program INTERNIST (v novější verzi nazývaný CADUCEUS), zaměřený na diagnostiku ve vnitřním lékařství. Odlišnost záleží nejen ve způsobu zpracování nejisté informace, ale převážně v řídicím algoritmu a s tím souvisejícím použitým formalismem pro reprezentaci znalostí. Důvod pro tuto odlišnost je v tom, že problémová oblast je zde nesrovnatelně rozsáhlejší, zatímco v případě programu MYCIN a jeho dosavadních aplikací jde převážně o relativně úzce vymezené oblasti.

Naším cílem v tomto článku není podat vyčerpávající přehled existujících expertních systémů a vzhledem k jejich množství a rychlosti vývoje si ani nemůžeme takový cíl klást. Pokusíme se spíše ukázat na ty rysy expertních systémů, které jsou většinou z nich společné. Čtenáře, který se zajímá o některé konkrétní systémy nebo o jejich použití, proto odkazujeme na [1] a zejména na sborníky z pravidelných mezinárodních konferencí IJCAI (International Joint Conference on Artificial Intelligence) a MEDINFO (Medical Informatics), kde se většina nově vznikajících systémů obvykle uvádí.

Abychom aspoň částečně navodili představu o současném stavu používání expertních systémů, uvedeme několik (více méně namátkou) vybraných systémů spolu s problémovými oblastmi, v nichž pracují.

Uvedli jsme již dva systémy určené pro medicínskou diagnostiku. PROSPECTOR se využívá v geologickém průzkumu, DENDRAL při interpretaci hmotových spektrogramů. Mikroprocesorová verze systému EXPERT slouží při vyhodnocování proteinové elektroforézy. Systém REX radí uživateli v používání metod regresní analýzy dat. SACON se osvědčuje při strukturální analýze mechanických komponent (při stavbě mostů, domů apod.), systém EL při návrhu a analýze elektrických obvodů.

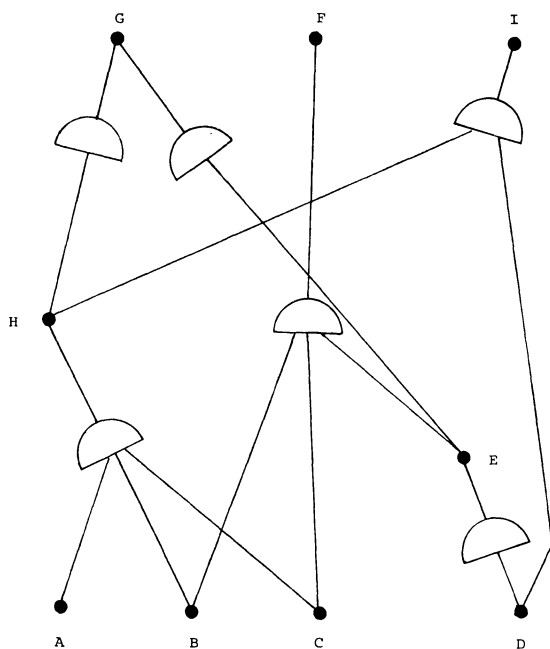
Protože aplikační oblasti expertních systémů jsou velmi rozmanité, můžeme nejprve na zjednodušeném příkladu ukázat základní myšlenky, na nichž jsou současné expertní systémy vystavěny, a v dalším naznačíme problémy, které je třeba řešit, chceme-li, aby expertní systémy dosáhly širšího praktického využití, tj. aby z laboratoří univerzitních a výzkumných pracovišť trvale postoupily do každodenní uživatelské praxe.

2. Příklad

Vyděme ze značně zjednodušené situace, kdy veškerá tvrzení relevantní pro danou expertní oblast mohou být vyjádřena seznamem výroků, např. A, B, C, D, E, F, G, H, I. Znalosti expertů budou vyjádřeny jako tato *produkční pravidla* [11]:

- | | |
|----------------------|--------------------------------|
| 1. $H \rightarrow G$ | 4. $H \& D \rightarrow I$ |
| 2. $D \rightarrow E$ | 5. $A \& B \& C \rightarrow H$ |
| 3. $E \rightarrow G$ | 6. $B \& C \& E \rightarrow F$ |

Jestliže nyní známe nějaká fakta o konkrétním zkoumaném případě, např. víme, že tvrzení A , B a C jsou pravdivá, můžeme na základě pravidla 5 uzavřít, že i H je pravdivé. Využijeme-li znalosti vyjádřené v pravidle 1, můžeme dále odvodit, že i G je pravdivé. Metodou přímého zřetězování (forward chaining) výroků jsme odvodili pravdivostní hodnoty některých dalších výroků. Čtenář se snadno přesvědčí, že kdybychom v našem případě znali ještě fakt, že D je pravdivé, mohli bychom pomocí našich znalostí v pravidlech 1 až 6 odvodit, že i všechny ostatní výroky jsou pravdivé, tj. že i E , F , I jsou pravdivé. Náznornou představu o tom, proč pro úplné ohodnocení všech výroků stačí přidat fakt, že D je pravdivé, lze získat tak, že reprezentujeme naše znalosti (pravidla) pomocí závislostní sítě, jak je uvedeno na obr. 1.



Obr. 1.



označení závislosti mezi antecedentem /konjunkcí podmínek/ a konsekventem pravidla.

V této reprezentaci je dobře vidět, že všechny výroky se rozpadají do tří skupin: Množiny vrcholových nebo též cílových výroků $\{G, F, I\}$, které se vyskytují pouze v závěru nějakého pravidla. Množiny mezilehlých výroků $\{E, H\}$, které se vyskytují jak v podmínkách, tak i v závěru nějakých pravidel, a množiny výchozích výroků (dotazů) $\{A, B, C, D\}$, které se vyskytují pouze v antecedentu nějakého pravidla a jejichž hodnoty určí uživatel. Touto sítí pravidel se můžeme pohybovat, jak je zřejmé, v zásadě dvojnásobným způsobem. Protisměrně od cílových výroků k podcílům až konečně k dotazům anebo od zodpovězených dotazů k jejich bezprostředním důsledkům získaným pomocí

pravidel zakódovaných v síti znalostí a konečně opakovaním postupu vytváření bezprostředních důsledků se můžeme dostat až k cílovým výročkům.

Čtenáři je již patrně zřejmé, že v tomto jednoduchém případě, kdy výroky nabývají dvou pravdivostních hodnot, vystačíme při vyhodnocování výroků s klasickým výrokovým počtem. Uvedený přístup však získá značně na zajímavosti, uvědomíme-li si, že ve většině obvyklých životních situací nejsme schopni zodpovídat dotazy s jistotou, ale umíme pouze vyjadřovat větší či menší míru přesvědčení o faktech. Typickým příkladem jsou výroky, s nimiž se setkáváme v medicínské diagnostice, kde expertní systémy našly dosud asi největší uplatnění. O leckterém pacientovi jsme ochotni říci, že je jistě obézní, ale v mnohých případech jsme s to vyjádřit pouze míru přesvědčení o jeho obezitě. Vyjadřujeme se tedy v termínech: spíše ano, spíše ne, nevím, skoro jistě apod. Zdrojem neurčitosti našich tvrzení může být jednak vágnost použitého predikátu, jednak neúplnost informace o třeba i rigorózním tvrzení.

Jedna z možností, jak se s tímto jevem vypořádat, je používat tzv. *faktoru věření* (jinými slovy zbohatit škálu pravdivostních hodnot), který připisujeme jednotlivým výročkům. Faktor věření (credibility factor) budeme brát např. z intervalu $\langle -1, 1 \rangle$ a hodnotu 1 budeme interpretovat jako „jistě ano“, hodnotu -1 jako „jistě ne“ a hodnotu 0 jako „nevím“, ostatní hodnoty jako větší či menší míru věření. Analogicky lze faktor věření přiřadit i samotným pravidlům. Faktor věření pravidlu je pak naše, přesněji expertova, důvěra v platnost pravidla v situaci, kdy předpokládám pravidla je přiřazen faktor věření 1. Terminologická poznámka: čtenáři možná cítí, že termín faktor věření příliš akcentuje subjektivní stránku věci. Proto se v literatuře někdy hovoří o stupních jistoty (certainty factor) [6], [15] nebo prostě jen o váhách, popřípadě se těchto termínů používá jako synonym [6]. My v dalším mezi nimi nebudeme rozlišovat. Požadavek na detailnější rozlišení vznikne v situaci, kdy se budeme zabývat teoretickým základem operací nad těmito vahami. Pak bude důležité, interpretujeme-li je např. algebraicky, pravděpodobnostně, logicky apod.

Náš příklad potom může vypadat třeba takto:

$$1'. H \rightarrow G (0.9)$$

$$2'. D \rightarrow E (0.9)$$

$$3'. E \rightarrow G (0.8)$$

$$4'. H \& D \rightarrow I (0.9)$$

$$5'. A \& B \& C \rightarrow (0.8)$$

$$6'. B \& C \& E \rightarrow F (0.9)$$

Budeme-li znát odpovědi na některé z dotazů, tj. budeme-li znát faktory věření pro dotazovací výroky, budeme moci odvozovat hodnoty faktorů věření cílových výroků. Nechť odpovědi na dotazy jsou třeba tyto faktory věření *cf*:

$$cf(A) = 0.7$$

$$cf(B) = 0.9$$

$$cf(C) = -0.9$$

$$cf(D) = 0.0,$$

což znamená, že o pravdivosti výroku *B* jsme přesvědčeni dosti silně, o pravdivosti

výroku A o něco méně, dosti silně jsme přesvědčeni o nepravdivosti výroku C a konečně o pravdivosti výroku D nevíme nic.

Když nyní budeme chtít na základě těchto nejistých odpovědí zjišťovat (aposteriorní) faktor věření pro cílové a mezilehlé výroky, musíme vyřešit nejprve tři problémy: a) definovat faktor věření pro složené výroky (v našem případě aspoň pro konjunkci); b) modifikovat faktor věření v závěr pravidla na základě váhy antecedentu a váhy pravidla. A konečně pro situaci, kdy několik pravidel vede k témuž závěru: c) určit, jak stanovit faktor věření tohoto závěru shrnutím „příspěvků“ jednotlivých pravidel.

Jedna z možných odpovědí na tyto otázky je tato:

a) Faktor věření pro složené výroky definujeme takto:

$$cf(A \& B) = \min(cf(A), cf(B))$$

$$cf(A \vee B) = \max(cf(A), cf(B)),$$

$$cf(-A) = -cf(A).$$

b) Jestliže pravidlo $A \rightarrow B$ má váhu $cf \geq 0$ a antecedent (výrok A) má faktor věření $a \geq 0$, pak závěr pravidla (výrok B) bude mít modifikovaný faktor věření $b = a \cdot cf$.

c) Jestliže dvě pravidla vedou ke stejnému závěru a přiřazují mu (každé samostatně) faktory věření cf_1 a cf_2 , pak operace, shrnující tyto „příspěvky“ a určující globální faktor jistoty závěru je definována výrazem

$$\frac{cf_1 + cf_2}{1 - cf_1 \cdot cf_2}.$$

Tato shrnovací funkce byla použita v systému PROSPECTOR. Není to však jediná vhodná shrnovací funkce. V systému MYCIN je shrnovací funkce definována takto:

$$cf_1 + cf_2 - cf_1 \cdot cf_2, \quad \text{když } cf_1, cf_2 \geq 0$$

$$\frac{cf_1 + cf_2}{1 - \min(|cf_1|, |cf_2|)}, \quad \text{když } cf_1 \cdot cf_2 < 0$$

$$cf_1 + cf_2 + cf_1 \cdot cf_2, \quad \text{když } cf_1, cf_2 < 0.$$

Lze ukázat, že existuje řada jiných funkcí, které splňují jisté jednoduché a přirozené požadavky a kterých je možno v této roli v podstatě rovnocenně použít. Které požadavky jsou přirozené, však může být předmětem diskuse. Podrobně se touto otázkou zabýval P. Hájek v [6].

3. Reprezentace znalostí

Uvedli jsme na začátku, že expertní systémy jsou programy založené na znalostech. V té souvislosti je ovšem možno namítnout, že vlastně v každém programu je skrytá jistá znalost o problému. Čím se tedy expertní systémy liší od jiných programů? Odlišnost záleží v tom, že programy založené na znalostech mají informaci o problémové oblasti a o tom, jak tuto informaci *explicitně vyjádřenou v bázi znalostí* používat. A to navíc v podobě, která je blízká struktuře znalostí expertů.

Čím víc znalostí, čím jsou kvalitnější, tím lépe. Každý expert, ale i každý program disponuje nějakými znalostmi a každý jich také nějak využívá. Lépe nebo hůře. Bohatých znalostí si vážíme, neméně si ovšem ceníme i umění dobře jich využívat pro řešení problémů. U lidských expertů mají znalosti velmi dynamickou strukturu. Znalosti v průběhu řešení nějakého problému *shromažďujeme, uchováváme, vybavujeme, měníme, odvozujeme a sdělujeme*. Tyto aktivity se odehrávají v paměti a je přitom často lhostejné, zda jde o paměť lidského nositele znalostí či o paměť počítače. Důležité je, že organizace znalostí v paměti může mít podstatný vliv na kvalitu a účinnost zmíněných aktivit.

Reprezentací znalostí v užším slova smyslu rozumíme zobrazení struktury problémového prostoru do sémantické paměti nositele znalostí. V širším slova smyslu reprezentace znalostí zahrnuje samo vymezení problémového prostoru včetně jazyka vyjadřujícího konceptuální entity, o něž nám jde. Není obtížné doložit, že vhodná reprezentace často sama přímo řeší danou třídu problémů, zatímco jiná naopak může řešení téměř znemožnit.

Názorným příkladem je třeba algoritmus násobení přirozených čísel, která jsou zapsána (reprezentována) jednou v arabské, podruhé v římské notaci. Zatímco v případě arabské notace se zmíněný algoritmus snadno naučí každý školák, při římském zápisu narážíme okamžitě na téměř nepřekonatelné potíže.

Klíčovým problémem znalostního inženýrství je tedy odpovědět na otázku „*Jak reprezentovat znalosti, které znalci (lidští experti) dané problémové oblasti očividně mají a používají*“.

To není jednoduché, protože expertům samým není mnohdy explicitní podoba jejich znalostí známa (neuvědomují si plně, co vědí), jejich rozhodovací postupy nejsou rigorózní, pravidla, jichž užívají, nejsou přesná. S tím vším se musejí tvůrci expertních systémů nějak vypořádat. Při vytváření bází znalostí pro specializované expertní systémy vznikají však nejen překážky teoretického charakteru, vyplývající z povahy věci, ale i překážky psychologické, vyplývající z povahových vlastností expertů a potenciálních uživatelů těchto systémů. Experti, od nichž jsou znalosti získávány, musejí být ochotni odhalit své nitro, zveřejnit, co vědí, i s rizikem, že bude odhaleno, co nevědí. Již dnes je možno pozorovat určitou nedůvěru nebo dokonce odpor k využívání expertních systémů. Důvody pro to jsou rozmanité. V některých zemích je možné pozorovat např. mezi lékaři odpor k používání medicínských diagnostických systémů z obavy, že je připraví o práci. Je to ovšem situace, která připomíná rozbíjení strojů v minulém století. Svůj podíl na odmítání expertních systémů má i určitá nedůvěra k výpočetní technice, zejména v zemích, které zaostávají za světovým vývojem ve spolehlivosti běžně dostupných počítačů. Problémy, které se v této souvislosti vynořují, však mohou mít i právní nebo i morální aspekty. Kdo např. bude odpovědný za stanovení diagnózy a terapie pacienta? Lékař, tvůrce expertního systému, nebo dokonce konstruktér počítače?

Existující expertní systémy se opírají převážně o *uniformní* prostředky reprezentace (tedy takové, kdy veškeré znalosti jsou vyjádřeny jednotným způsobem a uspořádány do jediné struktury). V tom je na jedné straně skryta vysoká účinnost programů ve vybraných, obvykle dosti úzkých problémových oblastech. Na druhé straně představuje uniformita dosti značné omezení pro rozsáhlé ambiciózní systémy budoucnosti, neboť ty by naopak měly umožňovat efektivní manipulaci s rozmanitými specializovanými znalostními strukturami patrně s možností *transformací a kompilace znalostí* pro účely kognitivní hospodárnosti.

Většina dosavadních přístupů má jedno společné. Stávají se neefektivní v souvislosti

s velkými znalostními (datovými) bázemi. Je tedy třeba věnovat velkou pozornost *strategiím* vyhledávání informace, tj. prostředkům, jak řídit aktivaci různých částí báze znalostí v situacích, kdy je víc možností a kdy tradiční selekční mechanismy selhávají.

Schank a Abelson [16] formulovali dva základní axiomy reprezentace znalostí takto:

1. Pro každé dvě položky báze znalostí, které mají stejný význam, by měla být jediná reprezentace v sémantické paměti.

2. Libovolná informace, která je implicitně obsažena v dané položce báze znalostí, musí být vyjádřitelná explicitně pomocí reprezentace významu této položky.

Tyto axiomy jsou, jak je patrné, motivovány požadavkem co nejvyšší hospodárnosti (časové i prostorové) při manipulacích se znalostmi. Samy však nejsou návodem pro případnou volbu vhodného formalismu.

Reprezentací báze znalostí z abstraktního hlediska se obvykle rozumí nějaká třída objektů spolu s relacemi a operacemi, jež slouží k výpovědím o nich a k manipulaci s nimi.

Báze znalostí je tedy z výpočtového hlediska *abstraktní datovou strukturou*. Abychom ji mohli efektivně používat, musíme být schopni ukládat informaci v bázi znalostí, dotazovat se na její jednotlivé položky, parcializovat svoje požadavky (např. „Najdi všechny objekty s takovou a takovou vlastností!“), přidávat a odstraňovat položky v bázi znalostí apod. Všechny tyto manipulace se znalostmi musejí být nezávislé na detailech konkrétní implementace. Důležitou operací s daty je například tzv. srovnávání (matching) dvou položek z báze znalostí. Při odvozování se pak výrazně uplatňuje proces *rozpoznávání*, v němž nové položky jsou srovnávány s množinami očekávaných prototypů. Cestu k těmto prototypům otevírají speciální strategie procesu inference, o nichž bude řeč v následujícím odstavci.

Je patrné, že prostředky reprezentace budou v podstatné míře ovlivněny cíli a předpokládanými způsoby použití navrhovaného expertního systému. Konkrétní volba prostředků bude záviset na *problémovém prostoru* (malý, velký, členěný, hierarchický apod.), na *datech a výchozích faktech* (spolehlivá, nespolehlivá, přesná, nepřesná, časově proměnná apod.) a na *inferenčních mechanismech* (uniformní, heterogenní, metaznalosti, potřeba kompilace znalostí, aktualizace inferenčních pravidel apod.).

Jako vyjadřovací prostředky jsou používány *formální jazyky* (logické kalkuly, gramatické systémy, např. hornovské klauzule apod.), *sémantické sítě a rámce* (plány, scénáře, skripty). Z hlediska typů inferenčních mechanismů se často používají *systémy, jejichž inference je vedena vzory* (pattern-directed inference systems) [18], *pravidly řízené systémy* (produkční systémy různých typů), *transformační systémy* (jednoprocesorové, produkční systémy založené na paralelních procesech apod.).

V souvislosti s problematikou reprezentace znalostí se často diskutuje dichotomie *procedurální* a *deklarativní* reprezentace. Zdá se, že je to dichotomie poněkud umělá, neboť reálné systémy v sobě zahrnují vždy oba aspekty. Deklarativní typ znalostí (typickým představitelem je predikátový kalkul a sémantická síť) zdůrazňuje fakta typu „co“. Procedurální reprezentace naopak vyzvedá znalosti typu „jak“. Pojem *rámce*, který v současné době patří mezi intenzívně zkoumané prostředky reprezentace znalostí, v sobě zahrnuje jak deklarativní, tak i procedurální aspekt. I když byl původně navržen

M. Minským [10] jako prostředek pro vizuální analýzu scény, dnes se již široce uplatňuje v nejrozmanitějších oblastech znalostního inženýrství. Obecná definice rámce není sice ještě ustálena, lze však již dobře určit některé charakteristické vlastnosti rámce jako abstraktní datové struktury. V každém případě by to měla být rekursivní struktura s identifikátorem a volným počtem rubrik (stránek nebo parametrů). Důležitou vlastností rámců je rovněž hierarchičnost dědičné struktury. Řada vlastností se v hierarchii rámců dědí, tj. přenáší se na podřízené rámce. Někdy se též mluví o kontextovém stromu rámců [9]. Vyplňováním hodnot parametrů vznikají *instance* rámců. Přitom hodnotami mohou být i funkce (procedury); tím je do statické hierarchie rámců inkorporována dynamická struktura. Na základě úvah o rámcích byly navrženy a realizovány programové prostředky pro reprezentaci znalostí jako KRL (Knowledge representation language) a FRL (Frame representation language). Většina rámcových struktur je implementována pomocí jazyka LISP nebo v některém z jeho dialektů.

4. Inferenční stroj

V předchozím odstavci jsme řekli, že čím více je znalostí a čím jsou kvalitnější, tím lépe. To je však jen jedna stránka věci. V tomto odstavci se budeme zabývat otázkou, jak se znalostmi zacházet. Ono totiž pouze *mít znalosti* je málo, musíme je též umět efektivně využívat. Jinak bychom se snadno ocitli v situaci mnicha, který sice měl mnoho knih (a možná velmi dobrých), ale jejich informační obsah mu byl na hony vzdálen. Musíme tedy umět rychle nalézat informaci relevantní pro řešení daného problému a dedukovat potřebné závěry.

Nejlépe prozkoumaným formálním prostředkem logické inference je predikátový kalkul. Ten má ovšem i svá omezení. Inferenční metody založené na využití predikátového počtu, zejména *rezoluční metoda*, a rovněž systémy pro řešení problémů jako STRIPS a PLANNER jsou v literatuře (např. [11]) dobře a přístupně popsány. Proto se v tomto odstavci zaměříme spíše na netradiční přístupy, zejména v souvislosti se zpracováním neexaktní a nejisté informace.

Neexaktní inference

V příkladu uvedeném na začátku tohoto výkladu jsme viděli *jednu* z možností, jak pracovat s nejistými znalostmi. Je zřejmé, že na strukturu CF faktorů věření je možno klást různé požadavky. Minimálním požadavkem je, aby to bylo úplné či částečné uspořádání (complete partial ordering). Důležité však je, že pravidla se chápou jako vícehodnotová inferenční pravidla sestávající ze dvou složek. Syntaktická složka se týká výroků či formulí samotných a odpovídá obvyklému chápání dedukčního pravidla v situaci, kdy jsou k dispozici pouze dvě pravdivostní hodnoty. Druhou složkou inferenčního pravidla pak jsou operace na faktorech věření, které určují faktor věření v závěr pravidla na základě vah jeho premis. Jestliže \wedge , \vee jsou svazové operace průseku a spojení a \otimes operace součinu reziduovaného svazu, definované na struktuře faktorů věření CF, pak nejdůležitější inferenční pravidla je možno vyjádřit např. v této podobě:

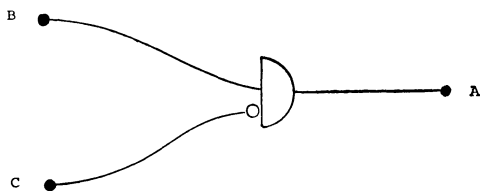
$$1. \frac{X[a], X[b]}{X[a \vee b]}, \quad 2. \frac{X[a], X \rightarrow Y[b]}{Y[a \otimes b]}.$$

První pravidlo říká, že když existuje argument pro věření ve výrok X s vahou a a současně argument pro věření v X s vahou b , pak můžeme odvodit X s vahou $a \vee b$. Druhé pravidlo je vlastně vícehodnotová analogie pravidla *modus ponens*. Podrobně se úvahami o neexaktní inferenci zabýval J. Pavelka v [12]. Obvykle se požaduje, aby operace na faktorech věření n -árního inferenčního pravidla byla izotónním zobrazením CF^n do CF , tj. zobrazením, které zachovává uspořádání faktorů věření.

Pravidla výše uvedených dvou typů budou zřejmě využita v každé bázi znalostí, která je reprezentována závislostní sítí, jak tomu bylo v našem příkladu. Operace spojení \vee odpovídá shrnování dílčích příspěvků věření v daný výrok; operace součinu \otimes odpovídá příspěvku věření odvozeném na základě jedné závislosti. V mnoha existujících expertních systémech [15], [19] je struktura faktorů věření interpretována pravděpodobnostně; obecný přístup reprezentovaný pracemi [6], [12] vyzvedá logický charakter inference a umožňuje začlenit pravděpodobnostní přístup pod tento jeden rámeček.

Nemonotónní inference

Dosud jsme uvažovali o znalostech jako o pravidlech typu „jestliže konjunkce podmínek, pak závěr“. Často ovšem potřebujeme používat méně rigorózních prostředků jako např. pravidel typu „ A vyplývá z B , ledaže by C “. Problematikou inference tohoto typu se v poslední době zabývá řada autorů [3], [8] mj. též v souvislosti s aktualizací bází znalostí. Systémy s pravidly typu „ A vyplývá z B , pokud lze konzistentně předpokládat C “ byly zkoumány v [14]. Inferenční systémy využívající pravidel uvedených typů jsou často nazývány nemonotónní systémy, protože přidání nové informace může vést k revizi platnosti již dříve odvozených tvrzení. Pravidla pro nemonotónní inferenci



Obr. 2. Schematické znázornění závislosti výroků A , B , C , které lze vyjádřit jako pravidlo typu: A vyplývá z B , ledaže by C . Kroužek na konci spojení vedoucího od výroku C naznačuje, že jde o nemonotónní vstup.

můžeme rovněž dobře znázornit pomocí grafové reprezentace. Na obr. 2 je znázorněna síť reprezentující závislost výroku A na výrocích B a C , kterou lze vyjádřit výše uvedeným nemonotónním pravidlem. Naším cílem je opět na základě pravdivostních hodnot některých výroků v síti znalostí odvozovat pravdivostní hodnoty ostatních výroků. K tomu je třeba stanovit, jak propagovat nejistou informaci nemonotónními pravidly. Jednou z možností je využít tzv. prahové logiky. Zajímavější je však definovat přímo nové operace na pravdivostních hodnotách (stupních jistoty), které by odpovídaly podmíněným závislostem.

Uvedme aspoň příklad, který naznačuje, jaké požadavky by tyto operace měly splňovat. Jestliže závislost (pravidlo) má stupeň jistoty 1 a všechny monotónní předpoklady jsou ohodnoceny hodnotou 1 a všechny nemonotónní předpoklady hodnotou -1 , pak závěr má stupeň jistoty 1.

Dosud jsme o znalostech uvažovali tak, že naše klasifikace výroků na cílové, mezilehlé a dotazovací byla jednou provždy daná a že pravidla (závislosti) svazující tyto výroky neobsahovala cykly. Znalosti lidských expertů však často mají tu vlastnost, že cykly obsahují a že některé výroky se mohou podle situace považovat někdy za dotazovací, někdy za cílové či mezilehlé. (Triviální příklad: „Jestliže pacient je muž, pak není žena“. „Jestliže pacient je žena, pak není muž“.) Existence cyklů sice působí jistě komplikace, ale z hlediska tvůrce expertního systému je spíše předností. Znalosti bez cyklů dovolují užívat jednoduchých inferenčních strategií, ovšem za cenu a priori daných cílů a dotazů. Chceme-li dosáhnout toho, aby na výroky, o něž nám jde, bylo možno se dívat různě podle situace, musíme zabezpečit, abychom v daném konkrétním chodu systému zabránili ohodnocování výroků v síti znalostí na základě cyklických důkazů. Musíme zabránit cyklům v uvažování, nikoli ve znalostech. Odkud budeme totiž čerpat informaci o výrocích, jejichž hodnoty nebudou již dále odvozovány (tj. které výroky budou dotazovací), může být silně ovlivněno konkrétní situací. Inference musí být dobře fundovaná, tj. algoritmus inference musí běžet po maximální podsíti znalostí neobsahující cykly (viz např. [3]).

5. Uživatel a expertní systém

Důležitým aspektem tvorby expertních systémů je způsob komunikace uživatele a systému. Není obtížné si uvědomit, že rozmanité formy informace mohou mít i zásadní vliv na komunikativní schopnosti dvojice uživatel – expertní systém. Důležité je předem vymezit ty situace, v nichž se bude expertní systém využívat. Mnoho situací bude patrně charakterizováno tím, že uživatelem nebude odborník ve výpočetní technice a skoro jistě ne v programování. Často se proto požaduje, aby informace mohla procházet expertním systémem na vstupu i výstupu v podobě vět přirozeného jazyka. Zde je rozsáhlé pole působnosti pro výpočtovou lingvistiku. Přirozený jazyk se může zdát být nejpřirozenější a též nejhodnější formou pro komunikaci (sdělování a přijímání) znalostí a faktů. Jsou ovšem i situace, kdy jsou vhodnější jiné přístupy jako např. zpracování obrazové informace apod. Při komunikaci pomocí přirozeného jazyka stojí před výpočtovou lingvistikou, ale i před logikou dosud jen v náznacích řešený problém, jak postihnout v „přirozené logice“ to, co se obvykle označuje termínem běžné uvažování (common sense reasoning). Může se zdát paradoxní, že se nám v umělé inteligenci daří poměrně úspěšně řešit problém automatizace uvažování vysoce kvalifikovaných odborníků ve speciálních oblastech, ale jen s obtížemi se daří postihnout uvažování lidí v nejběžnějších životních situacích.

Významným rysem expertních systémů je možnost vybavit je vysvětlovacími moduly. Dokonalý expertní systém by totiž měl nejen vydávat rozhodnutí, ale měl by být schopen poskytnout uživateli jednoduchým a srozumitelným způsobem informaci o tom, jak ke svým rozhodnutím dospěl. V systému MYCIN a v řadě dalších soudobých systémů se vysvětlení poskytuje (pokud vůbec) v termínech cílů a pravidel vedoucích k těmto

cílům. To je relativně úspěšný způsob při nepříliš rozsáhlých znalostních sítích a při poměrně krátkých inferencích. Dosud nedostatečně řešeným problémem je vysvětlování na základě kauzálních modelů a rovněž komprese detailních vysvětlení do vysvětlení globálních. Řešení takovýchto problémů nebo aspoň výstavba vhodného teoretického základu, na němž by se řešení mohlo hledat, je pro tvůrce expertních systémů stále otevřeným problémem.

Existence vysvětlovacích modulů expertních systémů má už dnes i přímé aplikace. Uspokojuje nejen přirozenou touhu člověka znát, jak a proč se dospělo k určitému řešení, ale může být využita i při výuce odborníků ve zvolené oblasti. Například diagnostický systém pro vyhodnocování rentgenových snímků lze využít jako výukový prostředek rentgenologů.

6. Závěr

Znalostní inženýrství a speciálně expertní systémy spolu se softwarovým inženýrstvím (programovou metodologií, relačními databázemi apod.) tvoří hlavní složky programového vybavení počítačů v japonském projektu počítačů 5. generace [5]. Rozvoj expertních systémů by měl vést ke kvalitativně nové symbióze člověka a počítače. To ovšem předpokládá intenzivní teoretický rozvoj a unifikaci technik znalostního inženýrství. Týká se to jak znalostí samotných a jejich reprezentace v počítači, tak i deduktivních neboli obecněji inferenčních postupů. V současné době je v projektu inteligentních počítačů silně akcentován logický rámec pro realizaci inferenčních postupů i pro reprezentování znalostí. Logické programování se v tomto projektu chápe jako významný nástroj syntézy znalostního i softwarového inženýrství spolu s databázovými systémy, architekturou počítačů a metodami umělé inteligence pro řešení problémů.

Literatura

- [1] A. BARR, P. R. COHEN, E. A. FEIGENBAUM (eds): *Handbook of artificial intelligence, Vol. 1–3*. Kaufman, Los Altos, 1981 a 1982.
- [2] R. DAVIS, D. B. LENAT: *Knowledge-based systems in artificial intelligence*. McGraw-Hill, 1982.
- [3] J. DOYLE: *A truth maintenance system*. *Artificial Intelligence* 12 (1979), 231–272.
- [4] R. O. DUDA, E. H. SHORTLIFFE: *Expert system research*. *Science* 222 (1983), 261–268.
- [5] K. FUCHI: *The direction the FGCS project will take*. *New Generation Computing* (1983), Vol. 1, 3–9.
- [6] P. HÁJEK: *Combining functions for certainty factors in consulting systems*. *Artificial Intelligence and Information-control Systems of Robots* (1982), 107–110.
- [7] F. HAYES-ROTH, D. A. WATERMAN, D. B. LENAT (eds): *Building expert systems*. Addison-Wesley, 1983.
- [8] D. McDERMOTT, J. DOYLE: *Non-monotonic logic I*. *Artificial Intelligence* 13 (1980), 41–72.
- [9] W. VAN MELLE ET AL.: *The Emycin manual*. Stanford Univ., Rept. No. STAN-CS-81-885, 1981.
- [10] M. MINSKY: *A framework for representing knowledge*. ve WINSTON (ed): *The psychology of computer vision*, McGraw-Hill, 1975.
- [11] N. J. WILSON: *Principles of artificial intelligence*. Tioga, 1980.

- [12] J. PAVELKA: *On fuzzy logic I — Many-valued rules of inference*. Zeitschr. f. math. Logik und Grundlagen d. Math. Bd. 25 (1979), 45—52.
- [13] P. RAULEFS: *Expert systems: State of the art and future prospects*. Informatik-Fachberichte 47, GWAI-81, Springer-Verlag, 1981, 98—111.
- [14] R. REITER: *A logic for default reasoning*. Artificial Intelligence 13 (1980), 81—132.
- [15] E. H. SHORTLIFFE: *Computer-based medical consultations: MYCIN*. American Elsevier, 1976.
- [16] R. C. SCHANK, R. P. ABELSON: *Scripts, plans, goals and understanding*. John Wiley & Sons, N. York, 1977.
- [17] M. STEFIK et al.: *The organization of expert systems, a tutorial*. Artificial Intelligence, North-Holland (1982), vol. 18, no. 2, 135—173.
- [18] D. A. WATERMAN, F. HAYES-ROTH (eds): *Pattern-directed inference systems*. Academic Press, 1978..
- [19] S. M. WEISS et al.: *A model-based method for computer-aided medical decision making*. Artificial Intelligence 11 (1978), 145—172.

Kvantová gravitace: sjednocený model reality?

P. C. W. Davies

Souhrn

Kvantové aspekty gravitace nabyly v poslední době velké astrofyzikální důležitosti. Hawkingův proces vypařování černých děr, tvorba částic v raném vesmíru a nejnověji inflační scénář raného vesmíru indikují, že kvantová gravitace mohla mít dominantní vliv na utváření struktury vesmíru ve velkém měřítku. Uvedené pokroky rovněž poukazují na možnost „samostatně se vytvořivšího vesmíru“, ve kterém se prostoročas i hmota objevují spontánně v důsledku kvantových efektů. To naznačuje, že mnohé z doposud uměle zaváděných počátečních podmínek, nutných ke sladění big bangového modelu s pozorovanými fakty, přirozeně a automaticky vyplývají z kvantové fyziky při velmi vysokých energiích (GUT a Planckova energie).

Z anglického originálu P. C. W. DAVIES: „*Quantum gravity: an unified model of existence?*“, uveřejněného v *Mitteilungen der Astronomischen Gesellschaft*, Nr. 58, Hamburg 1983, str. 47—56, s laskavým svolením autora i vydavatele přeložil ZDENĚK URBAN.

© 1983 Astronomische Gesellschaft, Hamburg.