Kapitola 2

Základní principy expertních systémů

Expertní poznatek:

Expertní systémy se poznají podle toho, že ze vstupního údaje "růže voní lépe než zelí" vyvodí závěr, že z růží bude také lepší polévka.

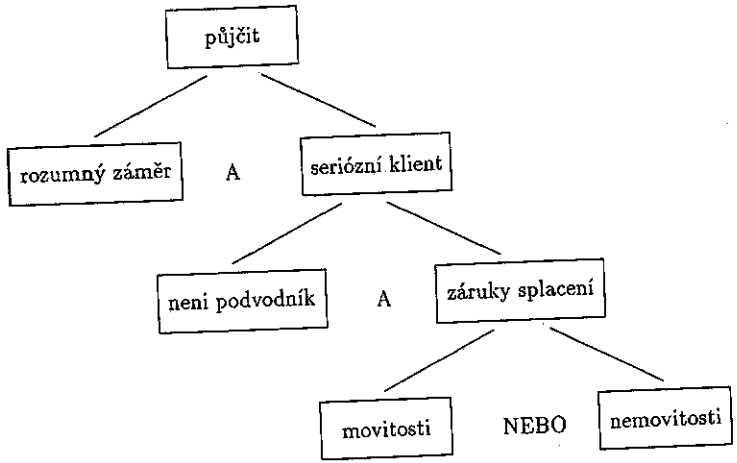
Murphyho počítačové zákony

2.1 Charakteristické rysy

Představme si BankuXY, která půjčuje peníze začínajícím podnikatelům. Zodpovědný pracovník banky nejprve posuzuje podnikatelský záměr. Zdá-li se mu tento záměr být rozumný, posuzuje seriózност klienta. Klient bude seriózní pokud to nebude žádný podvodník a pokud dá dostatečné záruky splacení úvěru. Záruky budou dostatečné, jestliže bude klient ručit movitým nebo nemovitým majetkem. Tyto *znalosti* lze schematicky znázornit v podobě *rozhodovacího stromu* (obrázek 2.1).

Podle E. Feigenbauma [19] je expertní systém *inteligentní počítačový program, který užívá znalosti a inferenční procedury k řešení problémů, které jsou natolik obtížné, že pro své řešení vyžadují významnou lidskou expertízu.* Znalosti nezbytné k činnosti na této úrovni plus použitá inferenční procedura mohou být chápány jako model expertízy nejlepších praktiků v oboru. Jiná možná definice je, že expertní systém je *počítačový program simulující rozhodovací činnost lidského experta při ře-*

Printed with FinePrint trial version - purchase at www.fineprint.com



Obr. 2.1: Způsob rozhodování o půjčkách

šení složitých úloh a využívající vhodně zakódovaných speciálních znalostí převzatých od experta s cílem dosáhnout ve zvolené problémové oblasti kvality rozhodování na úrovni experta [23].

V prvním období výzkumu v oblasti expertních systémů se kladl důraz na *kvalitu rozhodování na úrovni experta*; tomu ostatně napovídá už sám název *expertní* systémy. V pol. 70. let totiž převládal (optimistický) názor, že bude možné vytvořit "instantního experta", který by po nahrání na počítač byl schopen řešit i ty nejsložitější úlohy v dané oblasti. Tato představa narazila jak na meze samotných systémů (expertní systémy řešily úspěšně spíše rutinní problémy), tak i na skepsi uživatelů ("přece nebudu poslouchat nějaký počítač"). S jistou nedůvěrou uživatelů souvisejí i etické a právní otázky používání expertních systémů; kdo ponese odpovědnost za chybné rozhodnutí učiněné na základě doporučení expertního systému. Proto se postupně slevuje z pohledu na roli expertního systému v procesu rozhodování. Z původně zamýšlené (a deklarované) role *experta*, tedy někoho, kdo ví více než uživatel (a uživatel by tedy měl slepě poslechnout) se přechází k chápání expertního systému jako *kolegy* uživatele (tedy někoho, kdo ví přibližně stejně jako uživatel, ale

jehož výkonnost nepodléhá stresu, vlivům okolí apod.) nebo dokonce jako "pouhého" *asistenta* (tedy někoho, komu uživatel-expert svěřuje rutinní úlohy, aby se mohl plně soustředit pouze na komplikované případy). S tímto posuvem v pohledu na roli expertních systémů souvisí i posuv v terminologii. Dnes se spíše hovoří o *znalostních* systémech; důraz se tedy klade na *využívání vhodně zakódovaných speciálních znalostí převzatých od experta.*

Za charakteristické rysy expertních systémů se považuje:

- oddělení znalostí a mechanismu pro jejich využívání

  Znalosti experta jsou uloženy v bázi znalostí odděleně od inferenčního mechanismu. To umožňuje vytvářet problémově nezávislé (prázdné) expertní systémy (expert system shells), kde jeden inferenční mechanismus může pracovat s různými bázemi znalostí, obdobně, jako lze stejným způsobem místo "klíče k určování bonity klienta" procházet např. "klíč k určování hub" nebo "klíč k určování prince"[1].

- neurčitost v bázi znalostí

  V bázi znalostí jsou uloženy nejen exaktně dokázané znalosti, ale i nejrůznější heuristiky, které se např. expertovi osvědčily při rozhodování za dlouhou dobu jeho praxe. Zde se pak objevují pojmy jako "často" "většinou", které je potřeba kvantifikovat (např. v nějaké škále od "určitě ano" přes "nevím" až k "určitě ne"). Takovou znalostí s neurčitostí může například být "jestliže má pacient teplotu, obvykle je mu předepsán acylpyrin": pacient totiž "často" má teplotu v důsledku chřipky, ale "někdy" může mít teplotu, protože je v poúrazovém šoku.

- neurčitost v datech

  Konkrétní data o daném případu bývají zatížena neurčitostí způsobenou nepřesně určenými hodnotami nebo subjektivním pohledem uživatele (odpovědi na míru jistoty v nějakém tvrzení). Může to být například odpověď "snad ano" na dotaz, zda má pacient teplotu.

- dialogový režim

  Expertní systémy jsou nejčastěji konstruovány jako tzv. *konzultační systémy*. Uživatel komunikuje se systémem způsobem "dotaz systému - odpověď uživatele" obdobně, jako s lidským expertem. Tento systém práce byl do značné

[1] V případě této báze znalostí (která je někdy na přednášce používána jako ukázka) je adept princem, jestliže má malou nožku a je z dobré rodiny. Z dobré rodiny bude, když rodina má modrou krev a je bohatá; rodina je bohatá, když má peníze nebo pozemky. Tato báze znalostí má tedy stejnou strukturu jako báze pro BankuXY.

Printed with FinePrint trial version - purchase at www.fineprint.com

míry usnadněn nástupem osobních počítačů a s tím souvisejícím přechodem od dávkového zpracování na sálech výpočetních středisek k interaktivnímu zpracování přímo na psacím stole uživatele.

- vysvětlovací činnost

  Aby se zvýšila důvěra uživatelů v závěry a doporučení expertního systému, měl by systém poskytovat vysvětlení svého uvažování. Obvykle systém vysvětluje právě položený dotaz, znalosti relevantní k nějakému tvrzení, právě zkoumanou cílovou hypotézu, právě probíhající odvozování.

- modularita a transparentnost báze znalostí

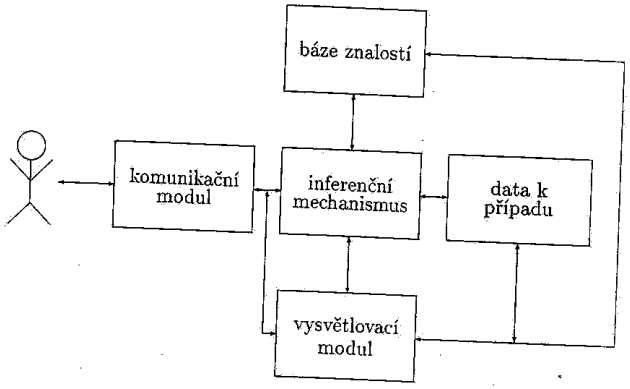
  Pro účinnost expertního systému je rozhodující kvalita báze znalostí. Modularita umožňuje snadnou aktualizaci báze znalostí, transparentnost umožňuje její snadnou čitelnost a srozumitelnost a kontrolu. Samotné vytváření báze znalostí probíhá iterativním způsobem při opakovaných konzultacích experta z dané problémové oblasti s odborníkem na tvorbu bází, tzv. znalostním inženýrem, kdy je báze znalostí postupně "laděna", až chování expertního systému (alespoň při konzultaci pro vzorové příklady) odpovídá představám experta.

Výše uvedené rysy odpovídají typickým expertním systémům (příkladem takového systému je např. MYCIN, (viz dále)). Některé systémy mohou pracovat v on-line režimu (tedy bez dialogu s uživatelem), jiné nemusejí využívat neurčitost. Nejpodstatnější charakteristikou je tedy oddělení báze znalostí a inferenčního mechanismu.

2.2 Struktura

Expertní systém je tvořen dvěma základními částmi: *bází znalostí* a *inferenčním mechanismem*. V bázi znalostí jsou uloženy znalosti experta z dané oblasti, inferenční (odvozovací) mechanismus umožňuje tyto znalosti využívat při konzultaci pro konkrétní případ. Dalšími komponentami jsou báze dat ke konzultovanému případu, vysvětlovací modul umožňující (do jisté míry) zdůvodnit postup systému při odvozování a modul pro komunikaci s uživatelem. Toto schéma (uvedené na obrázku 2.2) odpovídá "klasickému" systému, který pracuje v dialogu s uživatelem. Způsob komunikace s takovým expertním systémem pak obvykle vypadá tak, že:

- iniciativa je na straně expertního systému; ten volí otázky, které klade uživateli,
- uživatel se nespokojí pouze s doporučením systému, ale požaduje ještě jeho zdůvodnění.



Obr. 2.2: Obecné schema expertního systému

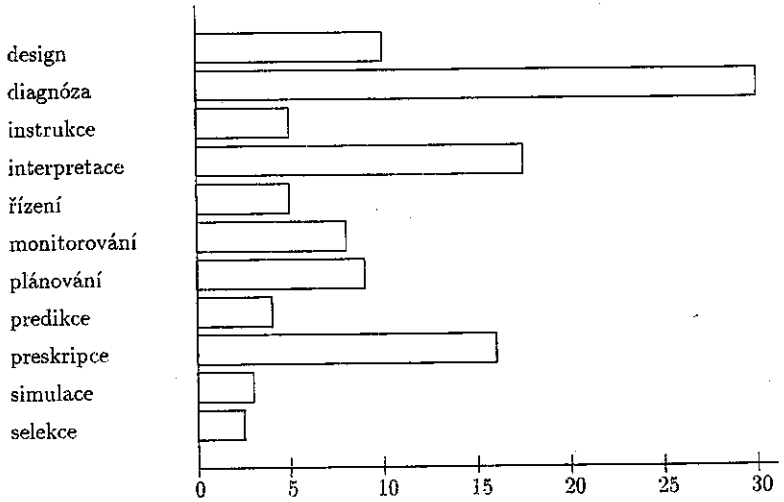
2.3 Typy expertních úloh

Podle charakteru řešených úloh se aplikace dají rozdělit [27] na:

- **diagnózu** - proces nalezení chyb či chybných funkcí systému (živého nebo neživého)
  např. MYCIN (viz dále) nebo INTERNIST - systém pro oblast vnitřního lékařství
  problémy: dysfunkce mohou být maskovány symptomy jiných dysfunkcí, dysfunkce mohou být jen občasné, některá data mohou být nepřístupná, chování systému nemusí být plně známo
- **interpretaci** - analýza dat s cílem určení jejich významu
  např. DENDRAL (viz dříve), PROSPECTOR (viz dále)
  problémy: data mohou být chybná, chybějící, cizorodá, zatížená šumem
- **monitorování** - průběžná (on-line) interpretace signálů a dat a určení okamžiku, kdy je nutná intervence
  např. VM - systém, který v reálném čase monitoroval pacienta napojeného na "umělé plíce"
  problémy: nebezpečí falešných alarmů

Printed with FinePrint trial version - purchase at www.fineprint.com

- **plánování** - nalezení posloupnosti akcí k dosažení cíle
  např. MOLGEN - systém pro plánování experimentů v molekulární genetice
  problémy: plánovací problémy jsou obvykle široké, komplikované a nedostatečně strukturované; ten kdo plánuje nemusí vždy chápat možné důsledky plánovaných akcí
- **návrh (design)** - vytváření konfigurací objektů vyhovujících daným podmínkám
  např. R1/XCON (viz dále)
  problémy: podobné jako při plánování, navíc hraje roli prostorové uspořádání
- **predikce** - předpověď běhu budoucích událostí na základě modelu minulosti a současnosti
  např. GLAUCOMA - systém pro predikci vývoje šedého zákalu
  problémy: predikce vyžaduje uvažování o čase a v čase



Obr. 2.3: Procento aplikací

Při hrubším dělení systémů vystačíme se dvěma základními skupinami; s expertními systémy *diagnostickými* (které řeší první tři typy úloh z výše uvedeného seznamu) a expertními systémy *generativními* (kam patří zbylé tři aplikace). Diagnostické expertní systémy pracují s pevným počtem cílů (diagnóz, hypotéz), ze kterých vybírají svá doporučení. Naproti tomu generativní expertní systémy si hypotézy generují dynamicky až v průběhu konzultace. Generativní systémy tedy na rozdíl od diagnostických mění "stav světa", se kterým pracují. Durkin [18] uvádí toto zastoupení jednotlivých typů úloh řešených expertními systémy (obrázek 2.3). Zatím převažují diagnostické systémy. Je to dáno jednak povahou problémů, kterou typicky řeší lidé - experti, jednak i tím, že vývoj diagnostické aplikace je relativně jednodušší.

V praxi je běžné, že aplikace může řešit více typů problémů. Již první expertní systém MYCIN (viz dále) nejprve diagnostikoval pacienta a pak předepisoval léky.

2.3.1 Diagnostické úlohy

Diagnostický proces (dle [51]) lze chápat jako získávání a interpretaci informací relevantních pro potvrzení přítomnosti nebo nepřítomnosti nějaké závady v systému. Empirický cyklus stanovení diagnózy je uveden na obrázku 2.4. Cyklus je tvořen třemi kroky: (1) formulování hypotézy, (2) testování hypotézy, a (3) přijetí nebo zamítnutí hypotézy.

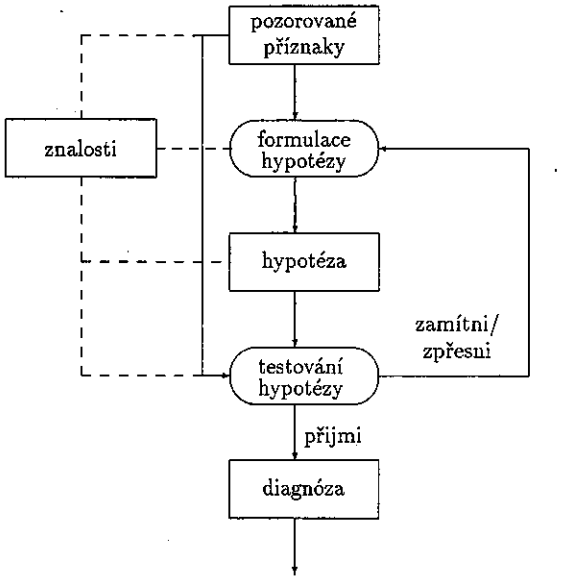
Konceptuální model diagnostické úlohy se opírá o možné typy znalostí použitých v dané aplikaci. Použité znalosti mohou mít charakter:

1. popisu *normálního* chování systému,
2. popisu *abnormálního* chování systému,
3. výčtu závad a seznamu příznaků pro každou závadu (bez explicitních znalostí o chování systému),
4. seznamu příznaků pro normální situaci.

V reálných aplikacích se výše uvedené znalosti různě prolínají, při uvažování různých konceptuálních modelů je ale vhodné mezi nimi rozlišovat (viz [51]):

- diagnostikování odchylky od normálu,
- diagnostikování porovnáním abnormálního chování,
- diagnostikování klasifikováním abnormality.

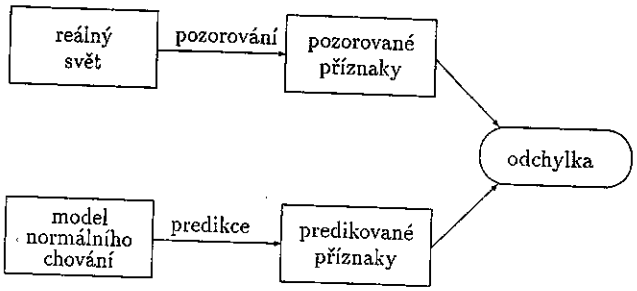
Printed with FinePrint trial version - purchase at www.fineprint.com



Obr. 2.4: Diagnostická úloha

Diagnostikování odchylky od normálu využívá první a čtvrtý typ znalostí. Tento způsob je vhodný v situacích, kdy není k dispozici dostatek znalostí o abnormálním chování a je známo pouze, jak vypadá chování normální (např. diagnostikování nového stroje). Při tomto způsobu diagnostikování se pozorované chování porovnává s chováním očekávaným a vyhodnocuje se případná odlišnost. Jak pozorované, tak očekávané chování bývá charakterizováno příslušnými příznaky (obr. 2.5).

Při *diagnostikování porovnáním abnormálního chování* se pracuje s modelem abnormálního chování systému. Využívá se tedy především druhý typ znalostí. Z předpokladu výskytu určité poruchy se dají predikovat některé abnormální příznaky. Ty se pak porovnávají s příznaky pozorovanými (obr. 2.6). Znalosti jsou obvykle formulovány jako kauzální vazby. Z důsledků (příznaků) se usuzuje na možné příčiny



Obr. 2.5: Diagnostikování odchylky od normálu

(poruchy), používá se tedy *abdukce*[2].

Oba výše uvedené způsoby, které využívají explicitní modely struktury a chování systému (a do jisté míry tedy simulují chování systému), bývají označovány jako systémy založené na modelech (model-based systems).

Obr. 2.6: Diagnostikování porovnáním abnormálního chování

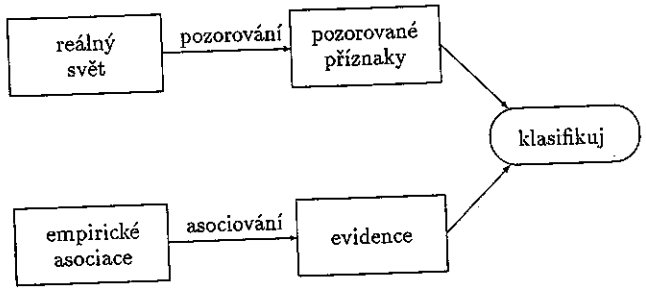
Třetí způsob, *diagnostikování klasifikováním abnormality* s modelem systému

[2]O způsobech odvozování se podrobněji píše v kapitole 4.1.

Printed with FinePrint trial version - purchase at www.fineprint.com

nepracuje (obr. 2.7). Využívá pouze více méně typické příznaky, které doprovázejí jednotlivé závady. Pracuje se tedy se třetím, někdy i čtvrtým typem znalostí. Použité znalosti se v tomto případě někdy označují jako "mělké" (shallow knowledge). Při řešení problému se zjišťuje, které závady jsou asociovány s pozorovanými příznaky[3]. Tyto znalosti jsou obvykle formulovány v podobě pravidel `z příznaku-x lze odvodit chorobu-y`. Příčinná souvislost (kausalita) je ale přesně opačná; `choroba-y způsobuje příznak-x`! Tomuto způsobu odvozování se někdy říká heuristická klasifikace[4], nebo `formuluj-hypotézu-a-testuj`. Jde o způsob typický pro diagnostické systémy první generace.

Na tomto místě jsme se zabývali pouze *statickým* aspektem stanovování diagnózy, tedy potřebnými znalostmi. Způsoby řešení diagnostické úlohy (tedy *dynamické* aspekty) jsou zmíněny v kapitole 4.1.

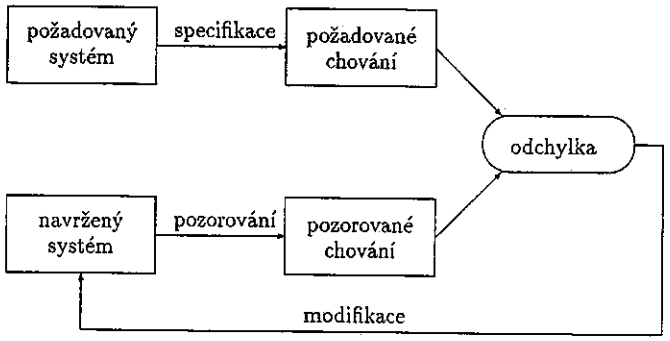


Obr. 2.7: Diagnostikování klasifikováním abnormality

2.3.2 Generativní úlohy

Plánování, tedy schopnost sestavit sekvenci akcí, která povede k dosažení požadovaného cíle, patří v umělé inteligenci k tradičním oblastem zájmu. Pohled na generativní úlohy (a plánování v AI obecně) se postupně vyvíjel. V prvním období bylo plánování chápáno jako dokazování teorémů nebo prohledávání (stavového prostoru nebo prostoru plánů).

[3]Někdy je použitý typ znalostí označován jako *empirické asociace*.
[4]Tento termín poprvé použil Clancey v roce 1985.



Obr. 2.8: Konfigurování

Dnes je plánování spíše chápáno jako splňování omezení, kladených na požadované řešení úlohy (constraint satisfaction). Požadavky na chování systému se porovnávají s chováním navrhovaného řešení. V případě, že řešení nevyhovuje, musí se modifikovat. Tímto způsobem se například řeší úlohy konfigurování nějakého systému (obr. 2.8). V situacích, kdy je k dispozici dostatek znalostí pro stanovení, co se má v dané chvíli provést, se postupuje způsobem `navrhni-a-aplikuj`. Klade se zde důraz na postupné rozšiřování dílčího řešení, a lze tedy postupovat poměrně přímočaře. V situacích, kdy dostupné znalosti jsou neúplné, nebo kdy existuje řada suboptimálních řešení, se strategie poněkud mění. Postupuje se způsobem `navrhni-a-reviduj`, který umožňuje navracení k předcházejícím variantám řešení.

Řešení je tedy u generativních úloh vytvářeno (poskládáno) z dílčích komponent. Tyto komponenty mohou být např. součásti nějakého zařízení (při designu) nebo dílčí akce (při tvorbě plánu).

2.3.3 Některé slavné systémy

MYCIN

V roce 1972 začal ve Stanfordu E. Shortliffe pracovat na programu, který by na základě rozboru krevních vzorků diagnostikoval infekční onemocnění a doporučoval léčbu vhodnými antibiotiky. Během dialogu s uživatelem-lékařem se systém ptá na zdravotní stav a zdravotní histo-

Printed with FinePrint trial version - purchase at www.fineprint.com