

Vážení zákazníci,

dovolujeme si Vás upozornit, že na tuto ukázkou knihy se vztahují autorská práva, tzv. copyright.

To znamená, že ukáзка má sloužit výhradně pro osobní potřebu potenciálního kupujícího (aby čtenář viděl, jakým způsobem je titul zpracován a mohl se také podle tohoto, jako jednoho z parametrů, rozhodnout, zda titul koupí či ne).

Z toho vyplývá, že není dovoleno tuto ukázkou jakýmkoliv způsobem dále šířit, veřejně či neveřejně např. umístováním na datová média, na jiné internetové stránky (ani prostřednictvím odkazů) apod.

redakce nakladatelství BEN – technická literatura
redakce@ben.cz



ČÁST A - MODELOVÁNÍ A UMĚLÁ INTELIGENCE

A.1 Z OBECNÉ TEORIE MODELOVÁNÍ

Modelování je proces tvorby modelů, v němž se v konvenčních přístupech využívá matematicko-fyzikální analýza i experimenty. *Model* je definován jako zobrazení existujících stránek reálného systému. *Identifikace* je pak proces ztotožnění modelu s objektem.

Identifikace a modelování jsou procesy, které se navzájem prolínají. Na základě apriorních informací o systému, determinovaném na zvolené rozlišovací úrovni a smyslu užití modelu přímo na objektu zkoumání, odhadneme pomocí matematicko-fyzikální analýzy *strukturu* modelu a vhodně volenými experimenty přímo na tomto systému pak v druhém kroku odhadujeme metodami identifikace hodnoty jeho *parametrů*.

Cílem identifikace a modelování je vytvořit takový model systému, jehož chování by bylo v jistém smyslu - nejčastěji z hlediska minima kritéria ztráty - stejné jako u systému za stejných provozních podmínek.

Zkoumaný systém nazýváme *procesem* a určený systém nazýváme *modelem*. Slovem *objekt* pak nazýváme onu hmotnou část skutečnosti, na které definujeme systém.

Definovat systém na objektu z hlediska daného účelu a vytvořit vyhovující model jsou základní úlohy identifikace a modelování.

Provedeme-li rozklad zkoumaného objektu na soubor složek a_1, a_2, \dots, a_k vzájemně vázaných a reprezentujících objekt, pak soubor složek

$$A = \{a_1, a_2, \dots, a_k\}$$

nazýváme *složkovou charakteristikou objektu* [1].

Označíme-li dále symboly r_{uv} vzájemnou souvislost složek a_u a a_v , tedy $r_{uv} = r(a_u, a_v)$, jako např. souvislost vstupních veličin složky a_u a výstupních veličin složky a_v , potom soubor všech uvažovaných závislostí

$$RC = (r_{uv}) \tag{1}$$

nazýváme *relační charakteristikou objektu*, a dále každý soubor

$$SYS = \{A, RC\}$$

nazýváme systémem a příslušný rozklad objektu představuje definování systému na objektu.

Uvažujme popis systému pomocí množiny vztahů

$$RC = \{RC_1, RC_2, \dots, RC_k\}$$

Vztah RC se určuje induktivním způsobem, v případě konvenčního modelování výběrem nějakého všeobecného vztahu a konkrétními hodnotami jeho neurčených složek, tedy

$$RC = \{F, \beta\}$$

kde F je struktura a β je množina neurčených složek (parametrů) vybraného vztahu.

V procesu identifikace pak vzhledem k definici systému na objektu rozeznáváme dvě etapy:

1. Výběr struktury systému F .
2. Porovnání chování systému se vztahy patřícími do množiny F s cílem nalezení optimálních hodnot složek vztahů (parametrů systému) β .

Druhou etapu (parametrickou identifikaci systému) je možno vykonávat systematicky, avšak za podmínky, že byla ukončena etapa první, t.j. výběr struktury (kterou vykonáváme heuristicky).

V každém okamžiku t z množiny uvažovaných časových okamžiků T ($t \in T$) působí na systém vstup $x(t)$, na nějž systém reaguje výstupem $y(t)$.

Z hlediska identifikace se jedná o stavbu takového funkcionálu, který přiřazuje pozorovaným vstupům $x(t)$ hodnoty pozorovaných výstupů $y(t)$

$$y(t) = F(x(t)).$$

Ve skutečnosti zde ještě existují nepozorované vstupy $v(t)$, které je třeba vzít v úvahu.

Úkolem identifikace je tedy vytvoření funkcionálu modelu F_M na základě pozorování vstupů x a výstupů y tak, aby F_M bylo blízké k F ve smyslu zvoleného kritéria.

V procesu identifikace se při tvorbě operátoru modelu můžeme opírat jak o informaci apriorní, tak o informaci aposteriorní. Apriorní máme k dispozici ještě před začátkem pozorování a aposteriorní nám přináší vhodně volený experiment.

Míra shody mezi reálným procesem a modelem je nejčastěji definována tak, že kritériem kvality bývá míra shody operátorů modelu F_M a operátorem procesu F , určovaná pomocí *funkce ztrát*

$$J(\mathbf{y}^0, \mathbf{y}^*) = (\mathbf{y}^0 - \mathbf{y}^*)^T (\mathbf{y}^0 - \mathbf{y}^*) = \mathbf{e}^T \mathbf{e}.$$

Taková kvadratická ztrátová funkce je funkcí skalární, která je funkcionálem výstupu procesu \mathbf{y}^0 a výstupu z modelu \mathbf{y}^* . Odchylka \mathbf{e} je mírou neshody mezi chováním procesu a modelem.

Optimálním výsledkem identifikace pak bude určení takového operátoru modelu F , pro který bude platit, že ztrátová funkce dosahuje minima

$$J(\mathbf{y}^0, \mathbf{y}^*)_{min}.$$

Konečným výsledkem identifikace by měl být takový operátor modelu F , který dává informaci o struktuře zkoumaného systému a o hodnotách parametrů matematického modelu.

Apriorní informace o zkoumaném procesu (jsou-li k dispozici) umožňují vytvořit výchozí struktury a proces identifikace pak spočívá v určení parametrů předpokládané struktury. Nejsou-li k dispozici apriorní znalosti o struktuře objektu, je třeba výchozí struktury rovněž identifikovat.

Experimentálním způsobem identifikace je řešení úlohy definice modelu procesu na základě souboru jeho vstupních a jim odpovídajících výstupních dat (datově orientované modely).

Je-li možno akceptovat předpoklad časové neproměnlivosti dynamických vlastností soustavy, je možno použít identifikaci jednorázovou. Není-li tento předpoklad splněn, je třeba použít metod identifikace průběžné, kdy parametry (případně i struktura) matematického modelu jsou vyhodnocovány průběžně v reálném čase.

A.2 MODEL Y KONVENČNÍ

Matematický model je vyjádřením problému pomocí fyzikálního a matematického formálního aparátu. Takto lze vyjadřovat složité vztahy symbolicky a zachovat při tom jednoduchost a racionálnost. Matematický model je představován soustavou matematických vztahů, jednoznačně popisujících zkoumaný jev nebo proces.

Formální aparát je tvořen obvykle rovnicemi algebraickými, obyčejnými nebo parciálními diferenciálními, soustavami takových rovnic, dále různými vztahy z teorie klasických množin, algebry, teorie pravděpodobnosti, matematické statistiky a matematické logiky. Tyto rovnice a vztahy se stávají matematickým modelem tehdy, jsou-li jednoznačně přiřazeny k určitému procesu nebo jevu.

Matematický model může vzniknout jako výsledek přímého zkoumání a pozorování jako datově orientovaný *model fenomenologický*, vyjadřující analyticky nebo numericky výsledky empirického zkoumání. Jinou cestou konstrukce konvenčního matematického modelu je dedukce, kdy model vzniká jako dílčí případ obecnějšího popisu. Tato cesta vede k *modelům asymptotickým*, vyjadřujícím jednoznačný matematický zákon chování systému nebo procesu, přičemž zákon je chápán jako dílčí případ obecné vědní teorie (Laplaceova rovnice je nejjednodušším případem obecné teorie fyzikálního pole).

A.3 MEZE APLIKACE KONVENČNÍCH MODELŮ

Konvenční matematicko-statistické analytické modely představují modely, pro jejichž sestavení je k dispozici (apriorně nebo aposteriorně) *přesná a úplná informace*. Pod přesnou a úplnou informací si zpravidla představujeme takovou informaci, která se dá reprezentovat či modelovat tak, že jak struktura tak i všechny parametry jsou jednoznačně určeny [2]. Modely, vykazující takovou formální dokonalost, nejsou zpravidla adekvátní skutečnosti, která je vágní a složitá.

Představa, že dostatečně složitý konvenční model může reprezentovat realitu s libovolnou přesností či adekvátností, není zřejmě správná. Formálně složité matematické modely vyžadují informace, které jsou náročné jak způsobem svého objektivního získávání, tak i nároky na svoji kvalitu. Tato skutečnost je zvláště závažná u modelů, určených pro práci v informačních nebo řídicích systémech reálného času. Potřebná rozsáhlá a náročná měření jsou v těžkých provozních podmínkách buď zcela nemožná, nadměrně náročná na údržbu, nebo při zajištění potřebné robustnosti je kvalita

jejich informace tak nízká, že jsou nepoužitelná. Na tuto skutečnost poprvé upozornil L. A. Zadeh [2], když v r. 1973 formuloval *princip inkompatibility* slovy:

„Tak, jak roste složitost nějakého systému, klesá naše schopnost činit precizní a přitom ještě použitelná tvrzení o jeho chování, dokud není dosaženo prahu, za níž se stávají preciznost a použitelnost (nebo relevantnost) téměř vzájemně se vylučujícími charakteristikami.“

Studovaná část reálného světa vykazuje zpravidla mnoho nejasného a vágního. Klasické metody pro formalizaci nepřesnosti předpokládají stochastický charakter ne zcela přesně determinovaných jevů. Takový stochastický přístup ke zpracování neručitosti pomocí aparátu pravděpodobnosti a matematické statistiky vyžaduje, aby příslušné jevy byly dobře definovanými prvky množiny a měly právě tak dobře definovaný význam výpovědí o nich. Dále je nutno dodržet řadu předpokladů o datech [3] a mít k dispozici dostatečný počet pozorování.

Inženýrské projekty řízení složitých technologických procesů ukázaly, že klasická matematická statistika se svým principiálním pojetím a řadou omezení není prostředkem k formalizaci a efektivnímu využití takového typu neurčitosti, který nazýváme *vágností* (pojmovou neurčitostí), jenž je při jazykovém popisu složitých systémů relevantní. V tomto směru dochází k jakési „krizi“ klasické numerické matematiky. Řešením této situace se zabývají následující kapitoly.

A.4 PROBLEMATIKA MODELOVÁNÍ SLOŽITÝCH SOUSTAV

Jistý rozpor mezi aplikovanou matematikou a reálným světem konstatoval L. A. Zadeh [2], když napsal: „Klasická matematika je příliš precizní při popisu systémů, kde je zahrnut lidský element“.

V kap. A.1 jsme definovali složkovou charakteristiku objektu vztahem (1)

$$RC = \{r_{uv}\}$$

kde

$$r_{uv} = r(a_u, a_v)$$

představuje souvislosti vstupních a výstupních veličin složek zkoumaného objektu.

Konvenční modely vycházejí z předpokladu, že relační charakteristika objektu je definována ostře, precizně a odchylky mezi odhadovanými a pozorovanými hodnotami závisle proměnné jsou tudíž výsledkem chyb pozorování. Původ odchylek mezi pozorovanými a estimovnými hodnotami závisle proměnné veličiny mohou však být nezanedbatelnou měrou způsobeny špatnou definovaností systémové struktury. Příčiny těchto odchylek můžeme hledat i v ne zcela ostrém charakteru systémových parametrů.

V této souvislosti vyvstává relevantní problém metod formalizace a efektivního zpracování neurčitých informací. Ukazuje se, že právě schopnost lidského mozku konstruovat a využívat jednoduché algoritmy pro vyvozování závěrů v podmínkách neurčitosti je hlavní příčinou kvality lidského uvažování.

Efektivní zpracování neurčitosti je v současné době již dobře rozpracováno [4]. Jsou definovány tři typy neurčitosti: fuzzitivnost (vágnost), nespécifičnost (špatné vymezení) a spor (konflikt), které jsou vyšetřovány v rámci pěti teorií, v nichž je vybudován aparát pro jejich kvantifikaci (teorie klasických množin, teorie fuzzy množin, teorie pravděpodobnosti, teorie možnosti a Dempster-Shaferova teorie). Přehled uvádí PŘÍL. 1.

Problematika poznávání obecných zákonitostí kognitivních procesů prostřednictvím modelování a simulací a snaha vytvořit metody a jim odpovídající systémy pro řešení složitých úloh takovými způsoby, které bychom považovali při řešení stejných úloh člověkem jako projevy jeho intelektu, je jednou z definic předmětu zájmu vědního oboru *umělá inteligence* [5]. Za jedny z jejích dosavadních výsledků můžeme považovat nekonvenční techniky modelování, simulací a řízení takových procesů, jejichž popis je vágní a pro jejich formalizaci nelze dobře použít klasickou metodu popisu neurčitosti - matematickou statistiku.

V řadě inženýrských aplikací, které využívají statistických metod pro vyjádření míry neurčitosti jevů, narážíme na problém malých rozsahů výběrových souborů. Počty vykonaných experimentů bývají často nedostatečné. Doplnková pozorování přitom často nejsou z technických nebo ekonomických důvodů dostupná.

Rozpor mezi informační náročností statistiky a omezenými informačními zdroji není jediný, na nějž se při použití klasických formálních prostředků na zpracování informací naráží. Další problém může vyplývat ze složitosti studovaných soustav a dějů, které v nich probíhají. Nemáme-li k dispozici dostatečný počet pozorování za reprodukovatelných podmínek, nelze variabilitu přírodních dějů eliminovat tím, že vyhovíme požadavkům reprodukovatelnosti stanovením příliš obecných podmínek. Získáme tím výsledky, jejichž rozptyl či konfidenční intervaly jsou natolik široké, že výsledky znehodnotí [3].

Potíže přetrvávají tehdy, pokud se snažíme pracovat pouze s *objektivně* zjištěnými informacemi. Problémy řeší metody, které připouštějí využití *subjektivně* zabarvených informací. Toho, že subjektivita je nedílnou součástí našeho poznání, si byl vědom již Bernoulli [6].

To jsou objektivní důvody, proč se ve zvýšené míře v řadě aplikací uplatňují metody umělé inteligence. Od ní se očekává, že nabídne východisko z problémů, které vznikly snahou o objektivizaci jak v přírodovědeckých, tak i inženýrských disciplínách.

A.5 VYUŽITÍ ZNALOSTÍ V PROCESU MODELOVÁNÍ

V souladu s rozvojem aplikací metod umělé inteligence vzniká trend přechodu od zpracování *údajů* ke zpracování *znalostí* [7]. Lidské znalosti můžeme rozdělit do dvou kategorií:

- a) V první jsou tzv. *znalosti hluboké*, jejichž zdrojem je vlastní poznání přírodních dějů ve formě přírodních zákonů. Jsou to znalosti objektivní, dostupné více či méně široké odborné veřejnosti. Jsou produktem analytických, abstraktních a teoretických postupů zkoumání jevů reality. Hluboké znalosti jsou vyjadřovány formou analytických matematických vztahů.
- b) Druhou kategorií znalostí označujeme jako *znalosti mělké*, povrchové (nikoli však povrchní !). Jsou to poznatky, jejichž zdrojem je dlouhodobá zkušenost, praxe a vlastní experimentování. Jsou to subjektivní znalosti, které kvalifikují úroveň experta. Mělké znalosti jsou vyjádřitelné formou predikátových kalkulů, heuristických pravidel, rámců či kvalitativních vztahů.

V procesu modelování hraje rovněž významnou roli dělení znalostí na znalosti *apriorní*, známé již ve fázi tvorby modelu a znalosti *aposteriorní*, získávané až v etapě jeho identifikace a využívání.

Využití hlubokých znalostí je typické pro metody konstrukce konvenčních, matematických statisticko-analytických modelů. Tyto modely, formalizované soustavami algebraických či diferenciálních rovnic, odrážejí díky aplikaci objektivních znalostí a metod obvykle širší třídy problémů a mají ponejvíce obecnou platnost. Uplatnění subjektivních znalostí v takových modelech není typické.

Snížená účinnost takových modelů v případech popisu složitých, špatně objektivně popsatečných a obtížně měřitelných soustav, vedla E. Feigenbauma [8] k vyslovení této téze:

„Přislíbem úspěšného řešení problémů jsou specifické, na danou problémovou oblast zaměřené znalosti (metody), kterými je třeba nahradit slabé metody všeobecné.“

Cesta uplatnění specifických (subjektivních, heuristických) znalostí vede k metodám tvorby *nekonvenčních nenumerních modelů*, které využívají vícehodnotové logiky a různé formální aparáty jak pro reprezentaci neurčitých pojmů, tak i pro aproximativní (přibližné) vyvozování.

Mělké znalosti vyjadřují pouze fenomenologické vztahy mezi entitami a nepodchycují tedy podstatu těchto vztahů. Nedostatečnost mělkých znalostí vede k odkazům na metaznalosti - znalosti o znalostech. Použití mělkých znalostí vede mnohdy ke zjednodušení principů a mechanismů metod vyvozování.

Přesto, že řada takto koncipovaných přístupů dosáhla komerčního úspěchu, v mnohých případech se projevila významnost limit jejich funkčních schopností.

V první řadě zde stojí skutečnost, že konstrukce a kvalita takového modelu je plně závislá na existenci kvalitního experta v dané problémové oblasti. Dále pak celou řadu problémů přináší role znalostního inženýra, jehož úkolem je vedení dialogu s expertem, čerpání jeho znalostí, volba vhodné metody jejich formalizace a způsobu aproximativního vyvozování. Vyvození, která takové systémy poskytují, mají vyhraněně lokální charakter.

Obecně je možno říci, že systémy, založené výhradně na mělkých znalostech, trpí ohraničenými možnostmi efektivního strukturování jimi realizovaných rozhodovacích procesů a zjednodušenými principy odvozovacích a vysvětlovacích mechanismů.

Snaha o přiměřené uplatnění hlubokých znalostí v nekonvenčních modelech z oblasti umělé inteligence vedla k přístupům, které umožňují integraci objektivních a subjektivních informací. Jejich výraznou charakteristikou je uplatnění globálních pohledů při rozhodování, plynoucích právě z uplatnění objektivních informací.

Jestliže typickým představitelem heuristického modelu jsou fuzzy modely, pak reprezentantem modelů, v nichž je možno vedle mělkých znalostí efektivně uplatnit i znalosti hluboké, jsou modely kvalitativní resp. semikvalitativní. Jejich význam podtrhuje skutečnost, že mohou dobře zahrnout i časové charakteristiky sledovaných dějů, a jsou tedy schopny modelovat soustavy dynamické.

Hluboké znalosti jsou rovněž nazývány znalostmi *kvantitativními*, mělké znalosti pak znalostmi *kvalitativními*. Možnost jejich integrace má zvláště velký význam v oblasti diagnostiky technologických procesů s ohledem na možnost predikce jejich poruchových stavů (bezpečnostní inženýrství).

A.6 INFORMAČNÍ ZTRÁTY PŘI MODELOVÁNÍ

Každé reálné zpracování reálných informací je nevratné a vede k informačním ztrátám. Z tohoto hlediska jsou výhodné takové metody modelování, které úpravy vstupních informací nevyžadují [9] nebo je minimalizují. Minimalizace úprav primárních dat a s tím spojená minimalizace ztrát informace je důležitá v takových oborech, jako je biotechnologie, spolehlivost, bezpečnost atd., kde je velký nedostatek primárních informací vůbec.

Označme množství informací v modelu I_m jako informační obsah modelu. Model musí splňovat nerovnost

$$I_p \geq I_m \quad (2)$$

kde I_p je potřebné množství informací, reprezentující požadovanou přesnost, obsažené v primárních datech. Označíme-li dI jako ztrátu informací při formálním zpracování, platí

$$I_p > I_m + dI \quad (3)$$

Další úvahy zjednodušíme předpokladem, že primární data jsou identická s výsledky experimentů. Tím eliminujeme subjektivně zbarvené znalosti.

Jelikož existuje přibližná funkční závislost mezi počtem optimálně rozmístěných experimentů a jejich informačním obsahem I_p , je možno informační obsah experimentů považovat za míru jejich počtu. Počet experimentů je shora omezen našimi možnostmi (technickými, ekonomickými, časovými apod.) na hodnotu I_M :

$$I_M > I_p \quad (4)$$

Proto pro malý počet experimentů, měřený dostupným informačním obsahem I_M ,

nemůže platit nerovnost

$$I_M > I_m + dI$$

kteřá ze vztahů (2) až (4) vyplývá.

Pokud z nějakého důvodu nelze snížit požadavek na výslednou přesnost modelu, je běžný postup v takovém případě takový, že se předmět studia počne modifikovat. Přistupuje se ke zjednodušení, např. k eliminaci vlivu některých nezávisle proměnných tím, že se považují za konstanty, případně jsou zanedbány.

Taková zanedbání nejsou učiněna na základě statistické analýzy, která by prokázala, že jsou statisticky nevýznamné. Takovou analýzu nelze provést pro nedostatek pozorování. Jsou zanedbány na základě úvahy podložené zkušeností, analogií nebo spekulací.

Mnohem vhodnější by bylo, pokusit se snížit ztrátu informací dI při formálním zpracování tím, že použijeme vhodnější formální aparát.

Vhodnost formálního aparátu může být hodnocena ze dvou hledisek: je to jednak skutečně menší informační ztráta, nebo schopnost zpracovat takové primární znalosti, které jsou pro klasický matematicko-statistický aparát nepřijatelné.

Označme množinu informací, nepřijatelných pro klasický model, jako NP . Musí platit

$$NP \in dI$$

Z tohoto hlediska se tedy NP nijak nepodílí na informačním obsahu modelu I_m .

Z hlediska potřebného množství informací I_p považujeme klasické matematicko-statistické modely za modely *informačně intenzivní*. Naproti tomu do skupiny modelů *informačně neintenzivních* řadíme modely nenumerické, jak budou postupně uvedeny dále.

Zajímavým problémem je jistě otázka, v jakých případech je vhodné použít ten který typ modelu. V kapitole L.4 je uveden přístup, naznačující postup, jímž lze vhodný typ modelu odhadnout na základě posouzení druhu a množství apriorních informací, které jsou dostupné pro jeho konstrukci, požadované přesnosti modelu a doby, která je pro jeho vytvoření k dispozici.

Kontrolní otázky k části A

1. Jaké jsou základní principy identifikace matematických modelů?
2. Čím jsou dány meze aplikace klasických matematických modelů?
3. V čem se odráží problematika modelování složitých soustav?
4. Čím se zabývá vědní obor Umělá inteligence?