

## 4 Reprezentace znalostí

V polovině 70. let se začal v umělé inteligenci přesouvat důraz od hledání univerzálního algoritmu pro řešení široké třídy úloh k práci se specializovanými znalostmi z určité oblasti. Tento trend našel své vyjádření v *expertních systémech*. Expertní systém můžeme chápat jako inteligentní počítačový program, který užívá znalosti a inferenční procedury k řešení problémů, které jsou natolik obtížné, že pro své řešení vyžadují významnou lidskou expertízu [Feigenbaum, 1979]. Znalosti nezbytné k činnosti na této úrovni plus použitá inferenční procedura mohou být chápány jako model expertízy nejlepších praktiků v oboru.

### 4.1 Základní pojmy

Znalost je lidský odhad uložený v mysli, získaný pomocí zkušeností a interakcí s okolním prostředím.

Znalost je fyzický, mentální nebo elektronický záznam o vztazích, o kterých věříme, že existují mezi skutečnými či imaginárními entitami, silami, jevy,

Znalost je vnitřní náhled, porozumění a praktické know-how, které všichni ovládáme – je to základní zdroj, který nám umožňuje chovat se inteligentně

Znalost je informace o světě, která umožňuje expertovi udělat rozhodnutí.

(Wikipedia)

Znalost je organizovaná informace využitelná na řešení problémů.

(Wolf 1990)

Znalost je informace, která je organizovaná a analyzovaná, aby se stala srozumitelnou a použitelnou pro řešení problémů nebo pro rozhodování.

(Turban 1992)

Znalost obsahuje pravdy a přesvědčení, perspektivy a koncepty, úsudky a očekávání, metodologie a know-how.

(Wiing 1993)

Znalost jsou fakta nebo ideje, které byly získány pozorováním, studiem, zkoumáním nebo zkušeností.

(Webster)

### **Znalosti:**

- **Explicitní:** formalizované, artikulované a tedy sdílené.
- **Implicitní:** primárně skryté (v datech) ale potenciálně formalizovatelné a tedy i sdílitelné.
- **Tacitní:** nevědomé a nesdílitelné znalosti skryté v myslích jedinců – expertů.

### **Znalosti:**

- **Deklarativní:** zachycující co platí (statické pravdy)
- **Procedurální:** zachycující jak postupovat při provádění nějakých akcí (usuzování)

### Požadavky na reprezentované znalosti

- Transparentnost
- Modulárnost
- Modifikovatelnost
- Užitečnost

## **4.2 Pohled znalostního managementu**

Anglický termín knowledge management je možno přeložit dvojnásobným způsobem: jako znalostní management nebo jako management znalostí. Tyto dva pojmy nelze zaměňovat. Zatímco management znalostí je „proces zachycení kolektivní expertízy organizace bez ohledu na to, kde se tato znalost nachází - v databázích, v dokumentech nebo v hlavách lidí, a její distribuce všude tam, kde pomůže dosáhnout nejvyšší přínosy.“ (Hibbard, 1997), znalostní management se týká způsobů využití znalostí při řízení organizace.

Znalostní management můžeme chápat jako praktickou odbornou činnost zaměřenou na využití znalostí v rozhodovacích a řídicích procesech za podpory informačních a komunikačních technologií. Znalostní management se tedy zabývá navrhováním, implementací a provozem systémů správy znalostí, jež zahrnují procesy získávání, reprezentace a zpracování, ukládání, vyhledávání a odvozování, prezentace, sdílení a distribuce znalostí. Teoretické zázemí tvoří kognitivní vědy a aplikační obory umělé inteligence (např. znalostní inženýrství), metody a techniky práce jsou odvozeny z praxe informačního managementu.

Znalostní management přichází s myšlenkou nabídnout správným lidem správné znalosti ve správný čas. Za tímto účelem se využívají nejrůznější technologie: intranet, systémy pro správu dokumentů (dokument management system), systémy pro správu obsahu (content management system), webové vyhledávače, push technologie, help-desk, groupware, workflow technologie, e-learning, znalostní systémy, datové sklady i metody dobývání znalostí z databází.

#### 4.2.1 Data, informace, znalosti .....z pohledu znalostního managementu

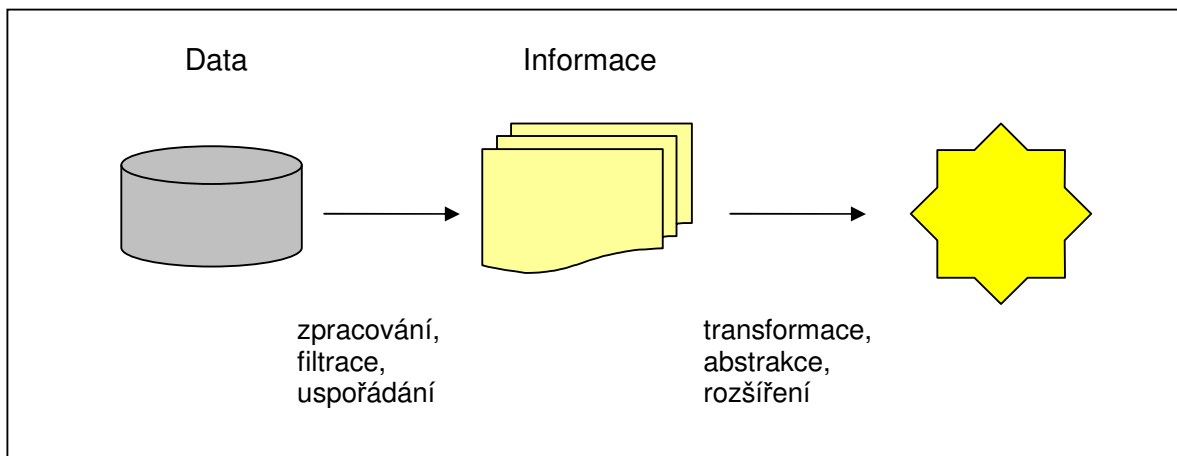
Zmiňme ve stručnosti některé pohledy na vztah mezi daty, informacemi a znalostmi, které byly formulovány v oblasti znalostního managementu. Dle Tobina [Tobin, 1996] můžeme definovat následující hierarchický vztah:

1. Data (+ relevance + účel =)
2. Informace (+ aplikace =)
3. Znalosti (+ intuice + zkušenosti =)
4. Moudrost

Podobnou hierarchie uvádí Beckman (Beckman, 1997):

1. Data (+ význam + struktura =)
  2. Informace (+ uvažování + abstrakce + aplikace =)
  3. Znalosti (+ výběr + zkušenosti + principy + ohraničení + učení =)
  4. Expertíza (jednotlivý expert) (+ integrace + distribuce + navigace =)
  5. Kompetence (způsobilost organizace)
- (Beckman, 1997)

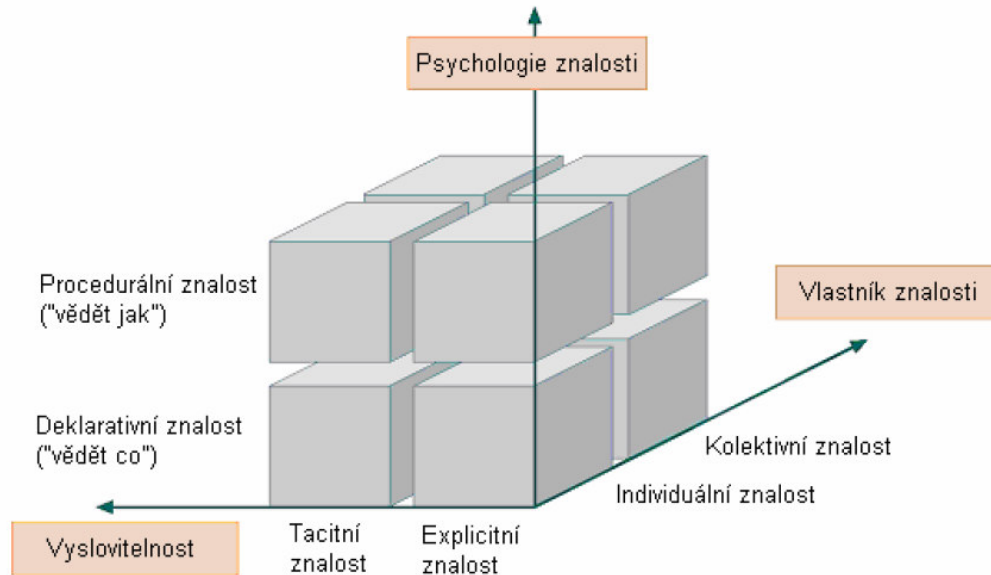
Z „technologického“ pohledu se na vztah mezi daty, informacemi a znalostmi můžeme dívat tak, jak je uvedeno na obr 5. [Mikulecký, 2006].



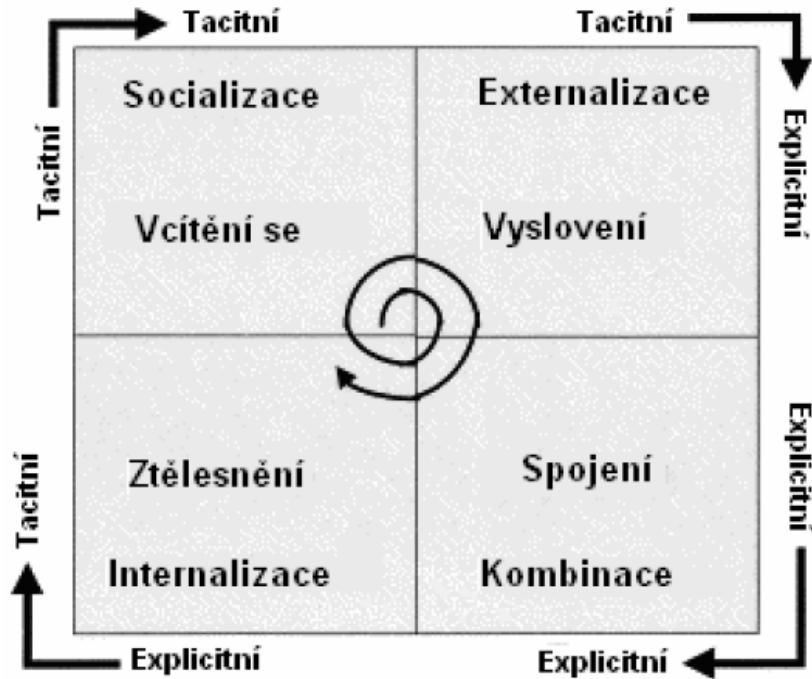
Obr. 4.1 Přechod od dat ke znalostem

Často používaná kategorizace znalostí využívá tři úhly pohledu (dimenze), podle kterých rozlišuje jednotlivé typy znalostí:

- psychologie znalostí (deklarativní vs. procedurální)
- vyjádřitelnost (explicitní vs. tacitní)
- vlastník znalostí (kolektivní a individuální)



Obr. 4.2 Tři dimenze znalostí (Bureš)



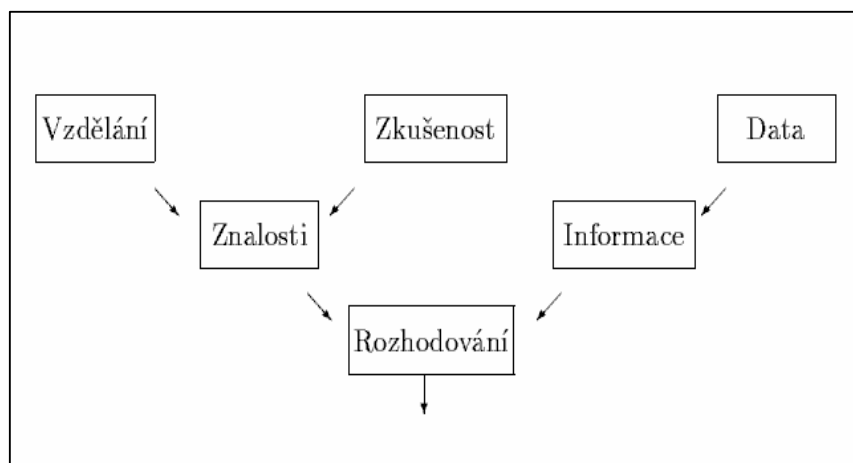
Obr. 4.3 Model SECI (Bureš)

Obr. 4.3 ukazuje tzv. SECI model životního cyklu znalostí formulovaný I. Nonakou v roce 1995. Tento model zachycuje 4 procesy tvorby znalostí - socializaci, externalizaci, kombinaci a internalizaci:

- socializace je proces přenosu nových nevyjádřených znalostí přes sdílenou zkušenost např. společným trávením času, formálními i neformálními schůzkami. Jde o získávání znalostí pozorováním, imitací či simulací
- externalizace je proces artikulace tacitních znalostí na znalosti explicitní. Neformulované znalosti se vyjadřují pomocí nějakého jazyka
- kombinace je proces přeměny jednodušších explicitních znalostí na komplexnější a systematictější
- internalizace je proces ztělesnění explicitních znalostí na znalosti tacitní např. v rámci různých školicích programů, studia dokumentů a metodik. Explicitní znalosti se v průběhu používání na základě zkušenosti a učení modifikují na znalosti neformulované.

### 4.3 Pohled znalostních systémů

#### 4.3.1 Data, informace, znalosti .....z pohledu procesu rozhodování



Obr. 4.4 Data a znalosti v procesu rozhodování (Widerhold)

Pojem znalost úzce souvisí s pojmy data a informace. Wiederhold [Widerhold, 1986] definuje znalosti a data z hlediska budoucích informačních systémů jako komplementární pojmy. Rozlišení dat a znalostí je podle něj následující:

- Jestliže se můžeme spolehnout při sběru materiálu na automatický proces nebo úředníka, hovoříme o datech. Správnost dat vzhledem k reálnému světu může být objektivně verifikována srovnáním s jeho opakovaným pozorováním.
- Jestliže hledáme experta, který by poskytl materiál, potom hovoříme o znalostech. Znalosti obsahují abstrakce a generalizace objemného materiálu. Obvykle jsou méně přesné a nemohou být objektivně verifikovány.

Data odrážejí současný stav světa na úrovni instancí, takže obsahují mnoho detailů, jsou objemná a často se rychle mění. Znalosti se tak často nemění. Pojednávají o generalizacích, o obecných zákonitostech. Rozdíl mezi daty a znalostmi se jeví podstatný v úvahách o aktualizaci, chybách v datech a nejistotě znalostí. Aplikace na podporu řízení a plánování budou záviset zejména na znalostech adekvátních pro rozhodovací postup a výjimky v datech budou spíše ignorovat.

Nově získaná data mohou být v rozporu s existujícími znalostmi. Ve velkých databázích se vždy vyskytnou chybné údaje, data, která jsou v rozporu se znalostmi, nemusí ihned vést k revizi znalostí. Jsou-li znalosti vybaveny faktory nejistoty, je třeba spíše tyto faktory za pomoci automatické učící se procedury aktualizovat. Někdy konflikt mezi daty a znalostmi ukazuje na neúplnost báze znalostí a lze jej vyřešit doplněním dalších znalostí, které pokrývají nové případy.

Jeden z problémů vztahu dat a znalostí spočívá v tom, že data na vyšší úrovni zpracování (např. již agregovaná nebo jinak odvozená) se mohou jevit jako znalost vůči nižší úrovni zpracování. To vede k úvahám o informačních systémech s mnohaúrovňovými strukturami.

Pojem znalosti je vhodné chápat v rámci logiky usuzování a inferenčního mechanismu, který ji prakticky realizuje. Za novou znalost považujeme tu informaci, která není daným inferenčním mechanismem odvoditelná z již dostupných znalostí (viz [Ivan2](#)). Ve znalostních systémech přitom můžeme pozorovat různé techniky odvozování:

- odvozování založené na případech: reprezentace znalostí získaných zkušeností (události, specifické případové studie a řešení).
- odvozování založené na pravidlech: znalosti přeložené do pravidel, která experti často používají při řešení složitých problémů
- odvozování založené na modelech: objektová technologie reprezentace a organizace znalostí, používající atributů, chování a vztahů objektů a simulaci procesů předmětné oblasti

#### 4.3.2 Role reprezentace znalostí

- Náhražka reality umožňující odhadovat důsledky našich akcí pouze na základě usuzování a nikoliv konání v reálném světě

Tato náhražka je nutně nepřesná a neúplná, při reprezentování znalostí se nevyhneme určitému zjednodušujícímu pohledu na realitu. Důsledkem je skutečnost, že občas dospějeme k chybným závěrům i tehdy, kdy používáme správné odvozovací postupy.

- Soubor ontologických závazků, tedy odpověď na otázku, v jakých pojmech máme usuzovat o světě

Volba základních pojmů výrazně ovlivní způsob reprezentování znalostí i odvozovací metody. Tak například medicínská diagnostika založená na pravidlech (např. expertní systém MYCIN) se podstatně liší od medicínské diagnostiky založené na rámcích (např. expertní systém INTERNIST). Zatímco MYCIN chápe medicínu jako soubor empirických asociací dávajících do vzájemné souvislosti symptomy a choroby, INTERNIST vidí soubor prototypů jednotlivých chorob uspořádaných do taxonomie, které se porovnávají s aktuálním případem. Od této základní volby se odvíjí řada dalších rozhodnutí. V případě

INTERNISTu je to např. podoba taxonomie, volba chorob, které budou reprezentovány, údaje o jednotlivých chorobách apod.

- Fragmentární teorie inteligentního usuzování

Návrh jednotlivých formalizmů reprezentování znalostí je motivován představou o jejich využití v průběhu usuzování. Tato představa je v jednotlivých formalizmech implicitně skryta – projevuje se jako odvozovací postupy, které jsou jednotlivými formalizmy reprezentace znalostí podporovány.

- Výpočetní prostředí ve kterém probíhá strojové usuzování

Čistě mechanisticky pojato, usuzování prováděné stroji (podle zastánců silné umělé inteligence i lidmi) je výpočetní proces. Otázka výpočetní efektivity tedy stojí v popředí při tvorbě reálné aplikace.

- Prostedí pro lidské vyjadřování, tedy jazyk, pomocí kterého mluvíme o světě

Použitý formalizmus musí být srozumitelný a snadno použitelný. Měl by umožňovat vyjádřit koncepty a pojmy z uvažované aplikační oblasti.

(Davis, Shrobe, Szolovits, 1993)

### 4.3.3 Znalostní modely a ontologie

V klasickém pojetí je získávání znalostí (např. pro expertní systémy) založeno na získávání znalostí od expertů. Zpočátku mělo získávání znalostí podobu transferu znalostí: znalostní inženýr přebíral znalosti od experta a přímo je vkládal do expertního systému. Takto vytvářené báze znalostí jsou ale obtížně modifikovatelné a přenositelné. Nebývají v nich totiž rozlišeny statické znalosti, týkající se celé aplikační oblasti a znalosti vztahující se k řešení dané konkrétní úlohy. Proto dochází na přelomu 80. a 90. let ke změně pohledu na proces získávání znalostí. Tento proces začíná být chápán jako **modelování znalostí**, tedy jako tvorba přehledných a opakovaně použitelných modelů dané úlohy. Znalosti jsou tedy zachycovány nezávisle na odvozovacích mechanismech a formalizmu reprezentace znalostí konkrétního expertního (znalostního) systému. Výhody tohoto přístupu jsou v zásadě dvojí:

- usnadnění vývoje aplikace: model vede tvůrce systému k lepšímu strukturování řešené úlohy,
- sdílení a opakované používání: pokud jsou modely založeny na standardizované terminologii, pak jsou srozumitelné nejen tvůrcům aplikace ale celé komunitě.

Nejnoveji se ve znalostním modelování objevuje pojem **ontologie**. Tento pojem je (na rozdíl od filosofického pojetí, kde ontologie znamená nauku o „bytí“) chápán jako označení domluvené terminologie pro určitou aplikační oblast, která umožňuje sdílení znalostí z této oblasti. Ontologie tedy umožňují formalizovat doménové znalosti:

1. Ontologie tvoří konceptuální popis znalostí – hraje roli meta-úrovně definující, co a v jaké podobě může být ve znalostech obsaženo,

2. Ontologie by měla být sdílitelná – neměla by být určena výhradně pro jedinou aplikaci. Předpokladem sdílitelnosti je přijetí daného způsobu konceptualizace v rámci širší komunity jako jistého standardu,
3. Ontologie je definovaná explicitně – nejde o ústní dohodu, ale o informace zachycené v určitém dokumentu pomocí jistého jazyka.

#### 4.3.4 Přístupy k reprezentování znalostí

Umělá inteligence přebírá představy o inteligentním usuzování z řady oblastí. Tyto oblasti rovněž daly vzniknout různým přístupům k reprezentování znalostí. Jednotlivé přístupy k reprezentování znalostí budeme ilustrovat na příkladu úlohy hodnocení bonity klienta banky, který žádá o úvěr. Jedná se vlastně o klasifikační úlohu, kdy na základě charakteristik klienta určíme, zda je bonitní či nikoliv.

Jeden pohled, historicky vzešlý z matematické logiky, vychází z toho, že inteligentní usuzování rovná se formální odvozování, typicky dedukce. Prostředky pro reprezentování znalostí zde vycházejí z výrokové resp. predikátové logiky a mají obvykle podobu pravidel.

If příjem(vysoký)	then úvěr(ano)	
If příjem(nízký) $\wedge$ konto(vysoké)	then úvěr(ano)	
If příjem(nízký) $\wedge$ konto(střední) $\wedge$ nezaměstnaný(ano)	then úvěr(ne)	
If příjem(nízký) $\wedge$ konto(střední) $\wedge$ nezaměstnaný(ne)	then úvěr(ano)	
If příjem(nízký) $\wedge$ konto(nízké)	then úvěr(ne)	

Obr. 4.5 Reprezentace znalostí pomocí pravidel

Druhý pohled, opírající se o psychologický přístup, vidí usuzování jako typicky lidské chování. Prostředky reprezentování znalostí nabízené tímto přístupem jsou rámce a sémantické sítě.

<p><b>Bonitní klient</b></p> <p>Is_a: klient  Příjem: vysoký nebo střední  Konto: vysoké nebo střední</p>	<p><b>NEBonitní klient</b></p> <p>Is_a: klient  Příjem: nízký  Konto: nízké  Majetek: žádný</p>
---	---

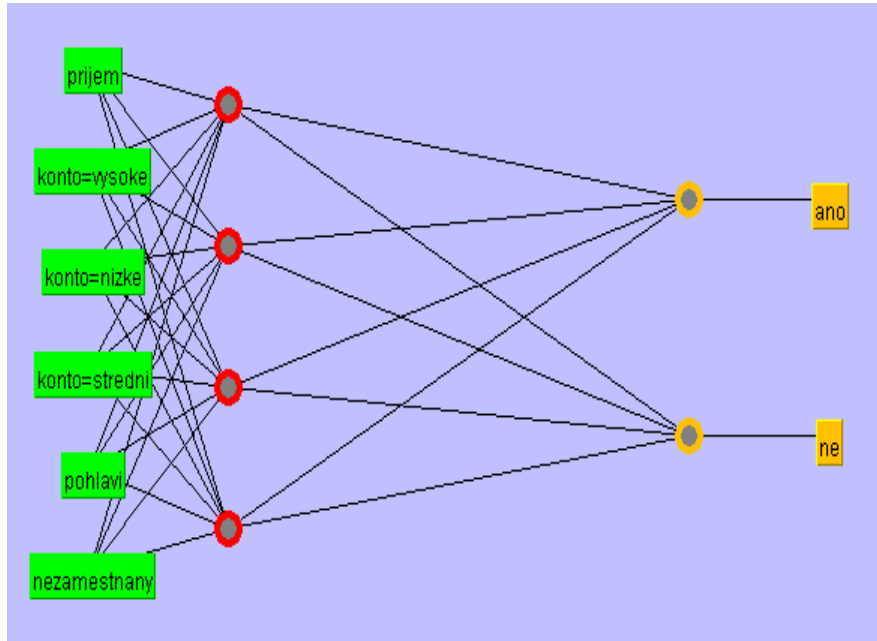
Obr. 4.6 Reprezentace znalostí pomocí rámců

Pohled vycházející z biologie vychází z toho, že klíčem k usuzování je architektura stroje založená na paralelním propojení velikého množství jednoduchých výpočetních jednotek. Nabízený formalismus jsou tedy neuronové sítě (Obr. 4.7).

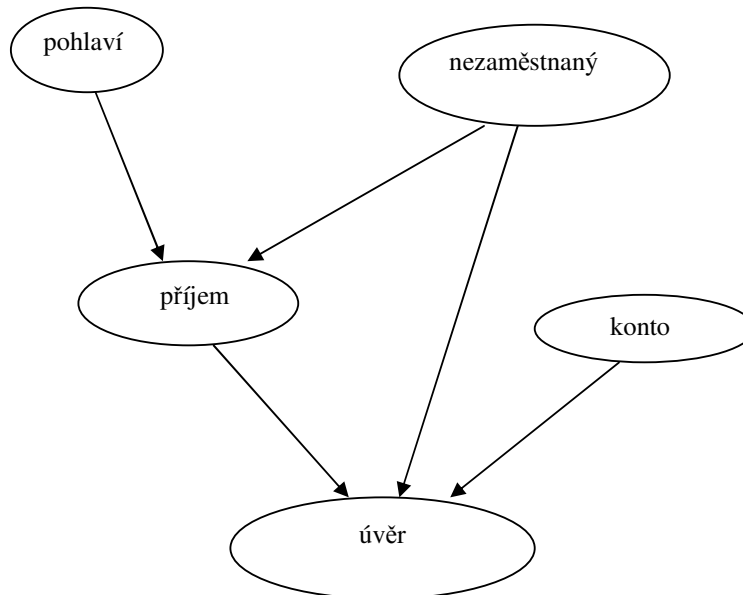


Pravděpodobnostní přístup propojuje logiku s neurčitostí. Na úrovni reprezentace znalostí našlo toto propojení svůj odraz v bayesovských sítích.

Pojetí vycházející z oblasti ekonomie se zaměřuje na otázku hodnot a preferencí. Je obvykle realizováno v systémech racionálních agentů.



Obr. 4.7 Reprezentace znalostí pomocí neuronové sítě



Obr. 4.8 Reprezentace znalostí pomocí bayesovské sítě

## 4.4 Výroková logika

Jazyk výrokové logiky je tvořen:

- výrokovými proměnnými ..... a,b,c,...,p,q,r
- logickými konstantami ..... T, F (někdy se píše 1, 0 což je označení, které budeme používat)
- logickými spojkami (v pořadí dle priorit) .....  $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$
- závorkami jakožto pomocnými symboly ..... ( )

Jazyk umožňuje vytvářet **formule**:

- výrokové proměnné a logické konstanty jsou formule (tzv. atomické)
- jsou-li  $\varphi$  a  $\psi$  formule, pak jsou formule i  $\neg\varphi, \varphi \wedge \psi, \varphi \vee \psi, \varphi \Rightarrow \psi, \varphi \Leftrightarrow \psi$

příkladem formule je tedy  $(\varphi \vee \neg\psi) \Rightarrow \neg\psi$

Zatím jsme se zajímali o syntaxi výrokové logiky. Sémantika přiřazuje formulím pravdivostní hodnotu na základě interpretace atomických formulí. Pravdivostní hodnoty logických spojek ukazuje následující tabulka

$\varphi$	$\psi$	$\neg\varphi$	$\varphi \wedge \psi$	$\varphi \vee \psi$	$\varphi \Rightarrow \psi$	$\varphi \Leftrightarrow \psi$
0	0	1	0	0	1	1
0	1	1	0	1	1	0
1	0	0	0	1	0	0
1	1	0	1	1	1	1

Tab. 4.1 Pravdivostní hodnoty základních logických spojek

## 4.5 Predikátová logika

Na rozdíl od predikátové logiky, kde jednotlivé výroky byly chápány jako dále nestrukturované, nyní nás bude zajímat vnitřní struktura tvrzení, se kterými budeme pracovat. Tomu odpovídá i použitý jazyk, tvořený:

- proměnnými a konstantami (pro pojmenování objektů světa, o kterém chceme vypovídat),
- predikátovými symboly (pro označení relací mezi objekty),
- funkčními symboly (pro označení funkcí),
- logickými spojkami .....  $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$
- kvantifikátory .....  $\forall, \exists$

Jazyk predikátové logiky opět umožňuje vytvářet formule, ale s vnitřní strukturou jednotlivých tvrzení. Základním výrazovým prostředkem predikátové logiky jsou **termy**. Termy jsou buď jednoduché (konstanty nebo proměnné), nebo složené (vzniklé aplikací funkce na termy, tedy např.  $\text{věk}(X)$ ). **Formule** pak vytváříme podobným způsobem jako ve výrokové logice:

- atomická formule má tvar  $P(t_1, t_2, \dots, t_n)$ , kde  $P$  je predikátový symbol a  $t_i$  jsou termy
- jsou-li  $\varphi$  a  $\psi$  formule, pak jsou formulemi i  $\neg\varphi, \varphi \wedge \psi, \varphi \vee \psi, \varphi \Rightarrow \psi, \varphi \Leftrightarrow \psi, \forall x \varphi(x), \exists x \varphi(x)$

Příkladem formulí jsou pak

$$\forall x \text{opice}(x) \Rightarrow \text{savec}(x)$$

$$(\forall \varepsilon > 0)(\exists \delta > 0) (\forall x) |x - a| < \varepsilon \Rightarrow |f(x) - f(a)| < \delta$$

Zatímco výroková logika je rozhodnutelná (o každé formuli lze na základě konečného počtu odvozovacích kroků říci, zda je logicky platná či není), predikátová logika rozhodnutelná není (platnost formule sice lze konečným počtem kroků potvrdit, ale neplatnost formule nemusí být ověřitelná v konečném počtu kroků). Proto se pro strojové usuzování znalostí reprezentují určitými podmnožinami predikátové logiky, které rozhodnutelné jsou.

### 4.5.1 Klauzulární logika

Klauzulární logika je podmnožina predikátové logiky, kde všechny formule jsou klauzule.

**Def: Literál** je atomická formule nebo její negace.

**Def. Klauzule** je formule, tvořená disjunkcí literálů.

$$L_1 \vee L_2 \vee \dots \vee L_n$$

Obecné klauzule mají podobu

$$\neg p_1 \vee \neg p_2 \vee \dots \vee \neg p_n \vee q_1 \vee q_2 \vee \dots \vee q_m$$

můžeme vyjádřit jako

$$p_1 \wedge p_2 \wedge \dots \wedge \neg p_n \Rightarrow q_1 \vee q_2 \vee \dots \vee q_m$$

**Def. Hornova klauzule** je klauzule obsahující nejvýše jeden pozitivní literál.

tedy např.

$$H \vee (L_1 \wedge \dots \wedge L_n)$$

kde  $L_i$  i  $H$  jsou pozitivní literály.

tuto klauzuli můžeme ekvivalentně zapsat jako implikaci

$$L_1 \wedge \dots \wedge L_n \wedge H.$$

**Věta:** Podmnožina predikátové logiky, kde formule jsou pouze Hornovy klauzule je rozhodnutelná.

Na Hornových klauzulích je založen programovací jazyk Prolog, navržený pro úlohy z oblasti umělé inteligence počátkem 70. let. Prolog pracuje s formulemi tří typů:

- atomické formule, které odpovídají složeným termům (lze je chápat jako fakta v databázi),

$$A_1$$

- podmíněné příkazy ve tvaru implikace  $A_1 \wedge A_2 \wedge \dots \wedge A_n \wedge B$   
 $B :- A_1, A_2, \dots, A_n$

- cílové klauzule, které mají tvar dotazu

$$?- C_1, C_2, \dots, C_n.$$

(kde atomické formule  $C_1, C_2, \dots, C_n$  se nazývají cíle)

Tento formalismus je postačující např. pro tvorbu expertních systémů, jejichž báze znalostí je tvořena pravidly.

## 4.5.2 Deskripční logika

Deskripční logika (DL) je pojem zastřešující celou řadu příbuzných logických formalismů. Jejich společným rysem je snaha zachytit strukturu tříd a relací chápaných obdobně, jako v systémech založených na rámcích resp. ontologiích.

Jazyk deskripční logiky je tvořen:

- konstantami a proměnnými
- unárními predikáty (nazývanými atomické koncepty)
- binárními predikáty (nazývanými atomické role)

Výrazy deskripční logiky se vytvářejí způsobem naznačeným v Tab. 4.2. Poznamenejme, že jednotlivé typy deskripční logiky se liší spektrem výrazů, které lze vytvářet. Tab. 4.2 tak odpovídá deskripční logice zvané  $\mathcal{ALCN}$ . Výrazy mají množinovou sémantiku nad universem instancí  $\Delta^I$ .

**Interpretační funkce**  $\cdot^I$  mapuje každý konceptuální výraz  $C$  na podmnožinu  $C^I$  universa  $\Delta^I$  a každý relační výraz  $R$  na binární relaci  $R^I$  nad kartézským součinem  $\Delta^I \times \Delta^I$ .

Výraz	syntaxe DL	(množinová) sémantika
atomický koncept	$A$	$A^I \subseteq \Delta^I$
atomická role	$R$	$R^I \subseteq \Delta^I \times \Delta^I$
top (nejobecnější třída)	$\top$	$\Delta^I$
bottom (nejspeciálnější třída)	$\perp$	$\emptyset$
negace (doplněk)	$\neg C$	$\Delta^I \setminus C^I$
konjunkce (průnik)	$C \sqcap D$	$C^I \cap D^I$
disjunkce (sjednocení)	$C \sqcup D$	$C^I \cup D^I$
univerzální omezení	$\forall R.C$	$\{x; \forall y (x,y) \in R^I \Rightarrow y \in C^I\}$
existenční omezení	$\exists R.C$	$\{x; \exists y (x,y) \in R^I \wedge y \in C^I\}$
omezení kardinality	$\geq n R$	$\{x;  y; (x,y) \in R^I  \geq n\}$
	$\leq n R$	$\{x;  y; (x,y) \in R^I  \leq n\}$

Tab. 4.2 Výrazy deskripční logiky ...

S využitím syntaxe predikátové logiky je možno výše uvedené výrazy zapsat následujícím způsobem:

Výraz	syntaxe predikátové logiky
atomický koncept	$A(x)$
atomická role	$R(x,y)$
top (nejobecnější třída)	$A(x) \vee \neg A(x)$
bottom (nejspeciálnější třída)	$A(x) \wedge \neg A(x)$
negace (doplněk)	$\neg C(x)$
konjunkce (průnik)	$C(x) \wedge D(x)$
disjunkce (sjednocení)	$C(x) \vee D(x)$
univerzální omezení	$\forall y; R(x,y) \Rightarrow C(y)$
existenční omezení	$\exists y; R(x,y) \wedge C(y)$
omezení kardinality	$\exists y_1..y_n; R(x,y_1) \wedge \dots \wedge R(x,y_n) \wedge \forall i \neq j y_i \neq y_j$ $\forall y_1..y_{n+1}; R(x,y_1) \wedge \dots \wedge R(x,y_{n+1}) \Rightarrow \exists i \neq j y_i = y_j$

Tab. 4.3 ... zapsané jako formule predikátové logiky

Báze znalostí je v deskripční logice tvořena terminologickými sentencemi (tzv. terminology box neboli TBox) používanými pro definování nových konceptů a relací mezi nimi, a přiřazovacími sentencemi (tzv. assertion box neboli ABox) pro přiřazování konstant proměnným (tedy pro unifikaci).

### TBox pro definování nových konceptů

$$A \equiv B$$

kde  $A$  je nový koncept a  $B$  je syntakticky správný výraz, např.

Woman	$\equiv$	Person $\sqcap$ Female
Man	$\equiv$	Person $\sqcap$ $\neg$ Woman
Mother	$\equiv$	Woman $\sqcap$ $\exists$ hasChild.Person
Father	$\equiv$	Man $\sqcap$ $\exists$ hasChild.Person
Parent	$\equiv$	Father $\sqcup$ Mother
Grandmother	$\equiv$	Mother $\sqcap$ $\exists$ hasChild.Parent
MotherWithManyChildren	$\equiv$	Mother $\sqcap$ $\geq 3$ hasChild
MotherWithoutDaughter	$\equiv$	Mother $\sqcap$ $\forall$ hasChild. $\neg$ Woman
Wife	$\equiv$	Woman $\sqcap$ $\exists$ hasHusband.Man

Obr. 4.9 Příklad ontologie definující rodinné vztahy

### TBox pro formulování axiomů o hierarchii konceptů

$$A \sqsubseteq B$$

kde  $A$  i  $B$  jsou koncepty pro které  $A^I \subseteq B^I$ , např.

$$\text{Man} \sqsubseteq \text{Person}$$

### ABox pro unifikaci

- pro koncepty

$$a : C$$

kde  $a^I \in C^I$ , např.

$$\text{jan} : \text{Man} \sqcap \forall \text{hasChild.Female}$$

- pro role

$$(a_1, a_2) : R$$

kde  $(a_1^I, a_2^I) \in R^I$ , např.

$$(\text{jan}, \text{eva}) : \text{hasChild}$$

## 4.6 Pravidla

Nejběžnější způsob reprezentování znalostí v expertních systémech je pomocí IF-THEN **pravidel**. Jde vlastně opět o reprezentaci znalostí založenou na matematické logice. Pravidla mohou být chápána dvojím způsobem; procedurálně:

JESTLIŽE situace PAK akce

nebo deklarativně:

JESTLIŽE předpoklad PAK závěr

kde situace, předpoklad a závěr jsou kombinace tvrzení o stavu světa. První (procedurální) interpretace je běžná v generativních expertních systémech; nastala-li příslušná situace, systém provede danou akci. Druhá interpretace odpovídá diagnostickým expertním systémům; je-li splněn příslušný předpoklad, systém odvodí daný závěr. Příkladem první interpretace je pravidlo z generativního systému R1/XCON (Obr. 4.10), příkladem druhé interpretace je pravidlo z diagnostického systému MYCIN (Obr. 4.11).

```
IF
  The current context is assigning devices to Unibus models AND
  There is an unassigned dual-port disk drive AND
  The type of controller it requires is known AND
  There are two such controllers, neither of which has any devices
    assigned to it, and
  The number of devices that these controllers can support is known
THEN
  Assign the disk drive to each of the controllers, AND
  Note that the two controllers have been associated and that each
    supports one drive
```

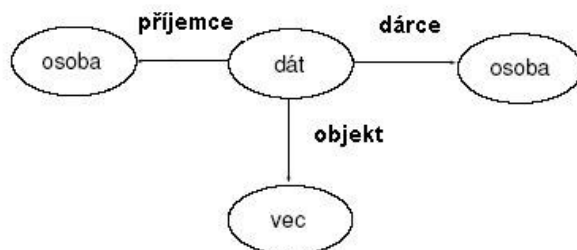
**Obr. 4.10 Pravidlo ze systému R1/XCON**

```
IF
  The site of the culture is blood, and
  The identity of the organism is not known with certainty, and
  The stain of the organism is gramneg, and
  The morfology of the organism is rod, and
  The patient has been seriously burned
THEN
  There is a weakly suggestive evidence (.4) that the identity
    of the organism is pseudomonas
```

**Obr. 4.11 Pravidlo ze systému MYCIN**

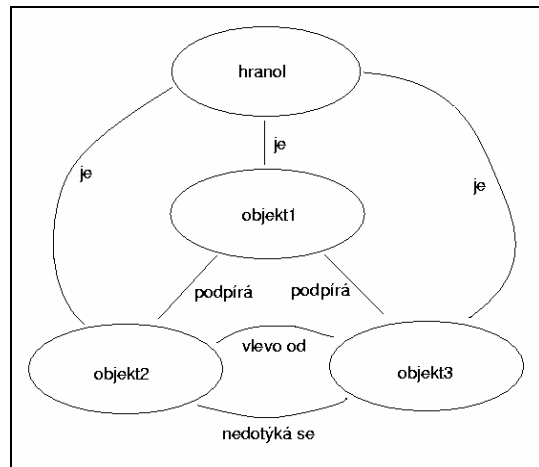
## 4.7 Sémantické sítě

Sémantické sítě byly navrženy R. Quillianem v druhé pol. 60. let v rámci prací na porozumění přirozenému jazyku jako model asociativní paměti člověka. Později byly zobecněny jako nástroj reprezentování znalostí v libovolné oblasti.



Obr. 4.12 Sémantická síť pro reprezentování významu slov přirozeného jazyka

Sémantická síť umožňuje popisovat realitu jako objekty, které jsou navzájem v nějakých vztazích (relacích). Sémantická síť má přirozenou grafovou reprezentaci; objekty jsou uzly a relace mezi nimi jsou hrany v grafu. Relace v sémantických sítích představují základní prostředek pro vyjadřování znalostí. Obr. 4.13 ukazuje příklad sémantické sítě, která definuje bránu složenou ze tří kostek.

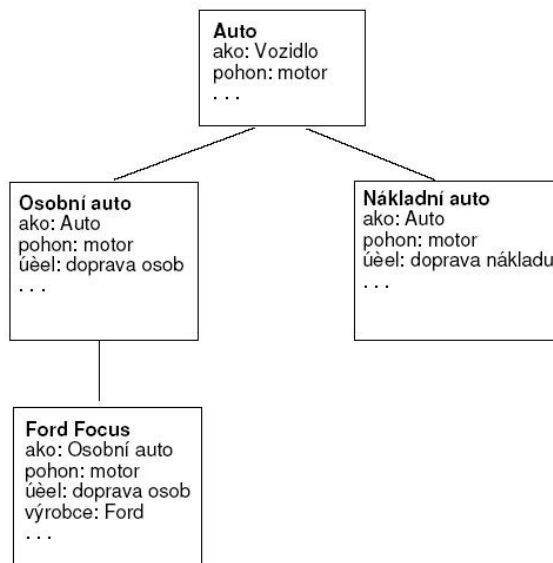


Obr. 4.13 Příklad sémantické sítě

## 4.8 1.6 Rámce

Rámce byly navrženy v polovině 70. let Marvinem Minským z MIT jako prostředek pro reprezentaci znalostí. Rámce v původní představě měly umožňovat reprezentovat stereotypní situace. Práce s rámci měla být založena na postupném vyplňování stránek, do kterých se zapisují hodnoty položek (vlastnosti). Přitom se hojně využívá předdefinovaných hodnot. Rámce dobře umožňují vyjádřit statické znalosti, tedy nějakou hierarchii pojmů (s použitím položky *a\_kind\_of*, zkráceně *ako*) nebo dekompozici (s použitím položky *part\_of*). Vazba mezi rámci se dá (podobně jako u sémantických sítí) znázornit grafem. Na rozdíl od sémantických sítí ale mají uzly v grafu

(rámce) vnitřní strukturu. V současné době rámce pronikly do programovacích jazyků. Zde se pro ně používá název objekty; příslušný styl programování využívající objekty se pak nazývá objektově orientované programování.



Obr. 4.14 Příklad hierarchie rámců

Formalizmus rámců umožňuje zachycovat znalosti v podobě tzv. případů, tedy typických úloh z dané aplikační oblasti, úspěšně řešených v minulosti.

## 4.9 Konceptuální svazy

**Svaz (lattice)** je matematické označení pro strukturu tvořenou množinou  $L$ , relací částečného uspořádání  $\leq$  (tato relace umožňuje definovat hierarchie) a dvěma operacemi  $\cap$  (největší dolní mez neboli infimum) a  $\cup$  (nejmenší horní mez neboli supremum). Obecně pro tyto operace platí:

- $a \cap b \leq a$ ,  $a \cap b \leq b$
- pokud  $c \leq a$  a  $c \leq b$ , pak  $c \leq a \cap b$
- $a \leq a \cup b$ ,  $b \leq a \cup b$
- pokud  $a \leq c$  a  $b \leq c$ , pak  $a \cup b \leq c$

**pozn:** Použití symbolů  $\cap$  a  $\cup$  není zcela náhodné. Množina všech podmnožin množiny  $U$  spolu s operacemi průnik ( $\cap$ ) a sjednocení ( $\cup$ ) a s relací inkluze ( $\subseteq$ ) tvoří svaz.

V konečných svazech bývá (z hlediska relace  $\leq$ ) definován maximální prvek  $\top$  a minimální prvek  $\perp$  (pro každé  $a \in L$  platí  $\perp \leq a \leq \top$ ).

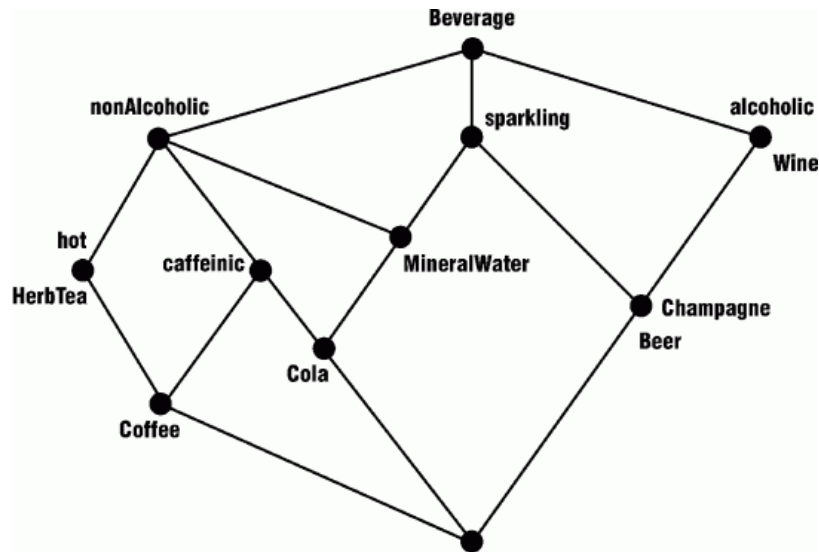


Na základě tzv. formální konceptuální analýzy můžeme data vyjádřená v podobě tabulky (např. Tab. 4.4) vyjádřit pomocí konceptuálního svazu, který dá do souvislosti objekty a atributy na základě relace  $\leq$  (Obr. 4.15). V tomto grafu konceptuálního svazu si povšimněme, že pivo a šampaňské je považováno za totožné koncepty. To je dáno skutečností, že na základě atributů v tabulce jsou oba koncepty nerozlišitelné.

Konceptuální svazy se používají pro reprezentování ontologií, pro jejich revizi, zjemňování a sdílení.

Concept Types	Attributes				
	nonalcoholic	hot	alcoholic	caffeinic	sparkling
HerbTea	X	X			
Coffee	X	X		X	
MineralWater	X				X
Wine			X		
Beer			X		X
Cola	X			X	X
Champagne			X		X

Tab. 4.4 Typy nápojů (extenze) a atributy (intenze) – dle Sowa



Obr. 4.15 Konceptuální svaz nápojů a jejich atributů (dle Sowa)

## Cvičení:

1) Doplňte ontologii uvedenou na Obr. 4.9 o definici konceptů Grandfather, Male, MotherWithoutSon, Son, Daughter (pro poslední dva koncepty použijte tzv. inverzní role  $R^{-}$  definované jako  $\{(x,y); (y,x) \in R\}$ )

2) S využitím deskripční logiky definujte následující koncepty z akademického prostředí (v zadání je naznačena hierarchie jednotlivých pojmů, jsou uvedeny i dostupné relace)

Person		
Worker		
Faculty		
Professor		
AssistantProfessor		
AssociateProfessor		
FullProfessor		
VisitingProfessor		
Lecturer		
PostDoc		
Assistant		
ResearchAssistant		
TeachingAssistant		
AdministrativeStaff		
Director		
Chair {Professor}		
Dean {Professor}		
ClericalStaff		
SystemsStaff		
Student		
UndergraduateStudent		
GraduateStudent		
Organization		
Department		
School		
University		
ResearchGroup		
Institute		

Relation	Argument 1	Argument 2
teacherOf	Faculty	Course
teachingAssistantOf	TeachingAssistant	Course
takesCourse	Student	Course
head	Organization	Person
undergraduateDegreeFrom	Person	University
mastersDegreeFrom	Person	University
doctoralDegreeFrom	Person	University
advisor	Student	Professor
subOrganization	Organization	Organization
affiliatedOrganization	Organization	Organization
member	Organization	Person
alumnus	Organization	Person
affiliateOf	Organization	Person
orgPublication	Organization	Publication
researchInterest	Person	Research
researchProject	ResearchGroup	Research
listedCourse	Schedule	Course
tenured	Professor	.TRUTH

## Literatura:

1. Baader F., Calvanese D., McGuinness D., Nardi D., Patel-Schneider P.: The Description Logic Handbook. Cambridge University Press 2003.
2. Beckman, T. A., Methodology for Knowledge Management, IASTED AI and Software Computing Conference, Banff, Canada, 1997.
3. Berka P. a kol.: Expertní systémy. Skripta VŠE, 1998
4. Bureš V., Čech P.: Knowledge Management at the Organizational Level. In: Proc. Znalosti 2006,
5. Davis R., Shrobe H., Szolovits P.: What is a Knowledge Representation? *AI Magazine* 14(1), 17-33, 1993.
6. Feigenbaum E.A.: Feigenbaum, E.A.: Themes and Case Studies of Knowledge Engineering. In: Michie, D. (ed.): Expert Systems in the Micro-Electronic Age. Edinburgh University Press, 1979.
7. Hibbard J.: Knowledge Management - Knowing What We Know. *Information Week*, 653, October, 1997.
8. Ivánek J.: Ivánek, J. a kol.: Vybrané kapitoly z matematické informatiky a znalostního inženýrství. Skripta VŠE, 1990.
9. Jirků P.: Logika. Neformální výklad základů formální logiky. Skripta VŠE, Praha 1993
10. Kelemen J. a kol.: Pozvanie do znalostnej spoločnosti. Iura Edition, Bratislava 2007
11. Lukasová A.: Formální logika v umělé inteligenci. Computer Press, Brno 2003
12. Mikulecký, P.: Znalosti a jejich význam ve znalostí společnosti. Seminář Znalostný manažment. VŠM Bratislava, 3.4.2006
13. Paralič J.: Manažment znalostí. Podklady k přednášce Univerzita Košice.
14. Sowa J.: Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks/Cole Publishing Co., Pacific Grove, CA, 2000
15. Tobin D.: Transformational Learning: Renewing Your Company Through Knowledge and Skills. John Wiley & Sons, 1996
16. Widerhold, G.: Knowledge versus Data. In: (Brodie, Mylopoulos eds.) On Knowledge Base Management Systems. Integrating Artificial Intelligence and Database Technologies. Springer Verlag, 1986.