

Intenzionální sémantika, reprezentace znalostí

Aleš Horák

E-mail: hales@fi.muni.cz

http://nlp.fi.muni.cz/poc_lingv/

Obsah:

- Intenzionální sémantika
- Reprezentace znalostí

Obsah

1 Intenzionální sémantika

- Logická analýza přirozeného jazyka
- Nedostatečná expresivita PL1
- Extenzionalismus PL1
- Extenze a intenze
- Transparentní intenzionální logika

2 Reprezentace znalostí

- Reprezentace znalostí
- Logika a sémantické sítě
- Rámce
- Pravidlové systémy
- Nejistota a pravděpodobnost

Logická analýza přirozeného jazyka

logická analýza PJ – analýza **významu** výrazů (vět) PJ

přirozený **jazyk** = nástroj *pojmového* uchopení reality

pojem – kritéria/procedury umožňující identifikovat různé konkrétní a abstraktní objekty

např. “**planeta**” – třída nebeských těles s určitými charakteristikami – obíhá po oběžné dráze kolem slunce, není zdrojem světla, ...

- **pojem** \neq **výraz** – např. výrazy v různých jazycích často reprezentují stejný pojem (pojem(“prvočíslo”) \equiv pojem(“prime number”))
- **pojem** \neq **představa** – představa je *subjektivní*, pojem je *objektivní*
- **pojmy** mohou identifikovat různé objekty:
 - jedno individuum – individuální pojmy (např. Petr, Pegas, prezident ČR)
 - třídu objektů – vlastnost (např. červený, šelma, hora)
 - *n*-člennou relaci – vztah (např. otec (někoho), křivdit (někdo někomu))
 - pravdivostní hodnotu – propozice (např. v Brně prší)
 - funkcionální přiřazení – empirické funkce (např. rychlost)
 - číslo – (fyzikální) veličiny (např. rychlost světla)

Logická analýza přirozeného jazyka

logická analýza PJ – analýza **významu** výrazů (vět) PJ

přirozený **jazyk** = nástroj *pojmového* uchopení reality

pojem – kritéria/procedury umožňující identifikovat různé konkrétní a abstraktní objekty

např. “**planeta**” – třída nebeských těles s určitými charakteristikami – obíhá po oběžné dráze kolem slánce, není zdrojem světla, ...

- **pojem** \neq **výraz** – např. výrazy v různých jazycích často reprezentují stejný pojem (pojem(“prvočíslo”) \equiv pojem(“prime number”))
- **pojem** \neq **představa** – představa je *subjektivní*, pojem je *objektivní*
- **pojmy** mohou identifikovat různé objekty:
 - jedno individuum – individuální pojmy (např. Petr, Pegas, prezident ČR)
 - třídu objektů – vlastnost (např. červený, šelma, hora)
 - *n*-člennou relaci – vztah (např. otec (někoho), křivdit (někdo někomu))
 - pravdivostní hodnotu – propozice (např. v Brně prší)
 - funkcionální přiřazení – empirické funkce (např. rychlost)
 - číslo – (fyzikální) veličiny (např. rychlost světla)

Logická analýza přirozeného jazyka

logická analýza PJ – analýza **významu** výrazů (vět) PJ

přirozený **jazyk** = nástroj *pojmového* uchopení reality

pojem – kritéria/procedury umožňující identifikovat různé konkrétní a abstraktní objekty

např. “**planeta**” – třída nebeských těles s určitými charakteristikami – obíhá po oběžné dráze kolem slunce, není zdrojem světla, ...

- **pojem** \neq **výraz** – např. výrazy v různých jazycích často reprezentují stejný pojem (pojem(“prvočíslo”) \equiv pojem(“prime number”))
- **pojem** \neq **představa** – představa je *subjektivní*, pojem je *objektivní*
- pojmy mohou identifikovat různé objekty:
 - jedno individuum – **individuální pojmy** (např. Petr, Pegas, prezident ČR)
 - třídu objektů – **vlastnost** (např. červený, šelma, hora)
 - *n*-člennou relaci – **vztah** (např. otec (někoho), křivdit (někdo někomu))
 - pravdivostní hodnotu – **propozice** (např. v Brně prší)
 - funkcionální přiřazení – **empirické funkce** (např. rychlost)
 - číslo – (fyzikální) **veličiny** (např. rychlost světla)

Logická analýza přirozeného jazyka

logická analýza PJ – analýza **významu** výrazů (vět) PJ

přirozený **jazyk** = nástroj *pojmového* uchopení reality

pojem – kritéria/procedury umožňující identifikovat různé konkrétní a abstraktní objekty

např. “**planeta**” – třída nebeských těles s určitými charakteristikami – obíhá po oběžné dráze kolem slunce, není zdrojem světla, ...

- **pojem** \neq **výraz** – např. výrazy v různých jazycích často reprezentují stejný pojem (pojem(“prvočíslo”) \equiv pojem(“prime number”))
- **pojem** \neq **představa** – představa je *subjektivní*, pojem je **objektivní**
- pojmy mohou identifikovat různé objekty:
 - jedno individuum – **individuální pojmy** (např. Petr, Pegas, prezident ČR)
 - třídu objektů – **vlastnost** (např. červený, šelma, hora)
 - *n*-člennou relaci – **vztah** (např. otec (někoho), křivdit (někdo někomu))
 - pravdivostní hodnotu – **propozice** (např. v Brně prší)
 - funkcionální přiřazení – **empirické funkce** (např. rychlost)
 - číslo – (fyzikální) **veličiny** (např. rychlost světla)

Logická analýza přirozeného jazyka

logická analýza PJ – analýza **významu** výrazů (vět) PJ

přirozený **jazyk** = nástroj *pojmového* uchopení reality

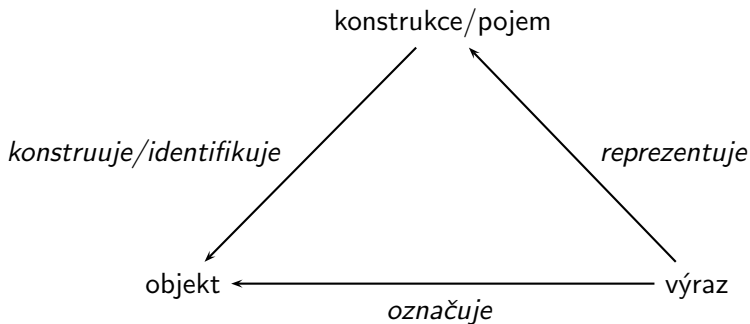
pojem – kritéria/procedury umožňující identifikovat různé konkrétní a abstraktní objekty

např. “**planeta**” – třída nebeských těles s určitými charakteristikami – obíhá po oběžné dráze kolem stálice, není zdrojem světla, ...

- **pojem** \neq **výraz** – např. výrazy v různých jazycích často reprezentují stejný pojem (pojem(“prvočíslo”) \equiv pojem(“prime number”))
- **pojem** \neq **představa** – představa je *subjektivní*, pojem je **objektivní**
- pojmy mohou identifikovat různé objekty:
 - jedno individuum – **individuální pojmy** (např. Petr, Pegas, prezident ČR)
 - třídu objektů – **vlastnost** (např. červený, šelma, hora)
 - *n*-člennou relaci – **vztah** (např. otec (někoho), křivdit (někdo někomu))
 - pravdivostní hodnotu – **propozice** (např. v Brně prší)
 - funkcionální přiřazení – **empirické funkce** (např. rychlost)
 - číslo – (fyzikální) **veličiny** (např. rychlost světla)

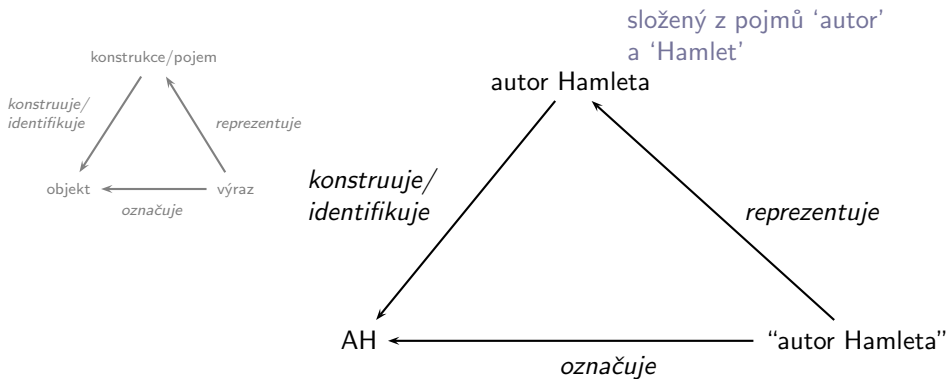
Vztah pojmu a výrazu

ve zjednodušené podobě: pojem odpovídá **logické konstrukci**



Vztah pojmu a výrazu

ve zjednodušené podobě: pojem odpovídá **logické konstrukci**



funkce ukazující v našem světě
na Williama Shakespeara

Omezenost predikátové logiky 1. řádu

dva omezující rysy:

- nedostatečná expresivita
- extenzionalismus

Expresivita: vyjadřovací síla jazyka

“Je-li barva stropu pokoje č. 3 uklidňující, je pokoj č. 3 vhodný pro pacienta X a není vhodný pro pacienta Y.”

analýza ve **výrokové logice**:

$P \Rightarrow (Q \wedge \neg R)$	P	“Barva stropu pokoje č. 3 je uklidňující.”
	Q	“Pokoj č. 3 je vhodný pro pacienta X.”
	R	“Pokoj č. 3 není vhodný pro pacienta Y.”

analýza v **PL1**:

$U(B) \Rightarrow (V(P, X) \wedge \neg V(P, Y))$	U	třída uklidňujících objektů
	B	individuum ‘barva stropu pokoje č. 3’
	V	relace mezi individuy ‘být vhodný pro’
	P	individuum ‘pokoj č. 3’
	X, Y	individua ‘pacient X’ a ‘pacient Y’

Omezenost predikátové logiky 1. řádu

dva omezující rysy:

- nedostatečná expresivita
- extenzionalismus

Expresivita: vyjadřovací síla jazyka

“Je-li barva stropu pokoje č. 3 uklidňující, je pokoj č. 3 vhodný pro pacienta X a není vhodný pro pacienta Y.”

analýza ve **výrokové logice**:

$P \Rightarrow (Q \wedge \neg R)$	P	“Barva stropu pokoje č. 3 je uklidňující.”
	Q	“Pokoj č. 3 je vhodný pro pacienta X.”
	R	“Pokoj č. 3 není vhodný pro pacienta Y.”

analýza v **PL1**:

$U(B) \Rightarrow (V(P, X) \wedge \neg V(P, Y))$	U	třída uklidňujících objektů
	B	individuum ‘barva stropu pokoje č. 3’
	V	relace mezi individuy ‘být vhodný pro’
	P	individuum ‘pokoj č. 3’
	X, Y	individua ‘pacient X’ a ‘pacient Y’

Omezenost predikátové logiky 1. řádu

dva omezující rysy:

- nedostatečná expresivita
- extenzionalismus

Expresivita: vyjadřovací síla jazyka

“Je-li barva stropu pokoje č. 3 uklidňující, je pokoj č. 3 vhodný pro pacienta X a není vhodný pro pacienta Y.”

analýza ve **výrokové logice**:

$P \Rightarrow (Q \wedge \neg R)$	P	“Barva stropu pokoje č. 3 je uklidňující.”
	Q	“Pokoj č. 3 je vhodný pro pacienta X.”
	R	“Pokoj č. 3 není vhodný pro pacienta Y.”

analýza v **PL1**:

$U(B) \Rightarrow (V(P, X) \wedge \neg V(P, Y))$	U	třída uklidňujících objektů
	B	individuum ‘barva stropu pokoje č. 3’
	V	relace mezi individuy ‘být vhodný pro’
	P	individuum ‘pokoj č. 3’
	X, Y	individua ‘pacient X’ a ‘pacient Y’

Omezenost predikátové logiky 1. řádu

dva omezující rysy:

- nedostatečná expresivita
- extenzionalismus

Expresivita: vyjadřovací síla jazyka

“Je-li barva stropu pokoje č. 3 uklidňující, je pokoj č. 3 vhodný pro pacienta X a není vhodný pro pacienta Y.”

analýza ve **výrokové logice**:

$P \Rightarrow (Q \wedge \neg R)$	P	“Barva stropu pokoje č. 3 je uklidňující.”
	Q	“Pokoj č. 3 je vhodný pro pacienta X.”
	R	“Pokoj č. 3 není vhodný pro pacienta Y.”

analýza v **PL1**:

$U(B) \Rightarrow (V(P, X) \wedge \neg V(P, Y))$	U	třída uklidňujících objektů
	B	individuum ‘barva stropu pokoje č. 3’
	V	relace mezi individuy ‘být vhodný pro’
	P	individuum ‘pokoj č. 3’
	X, Y	individua ‘pacient X’ a ‘pacient Y’

Nedostatečná expresivita PL1 – pokrač.

Červená barva je krásnější než hnědá barva. Kostka je červená.

analýza v PL1:

$Kr(\check{C}_1, H)$

$\check{C}_2(Ko)$

\check{C}_1 individuum 'červená barva'

\check{C}_2 vlastnost individuí 'být červený' (třída červených objektů)

nelze vyjádřit $\check{C}_1 \equiv \check{C}_2$

Nedostatečná expresivita PL1 – pokrač.

Červená barva je krásnější než hnědá barva. Kostka je červená.

analýza v PL1:

$Kr(\check{C}_1, H)$

$\check{C}_2(Ko)$

\check{C}_1 individuum 'červená barva'

\check{C}_2 vlastnost individuí 'být červený' (třída červených objektů)

nelze vyjádřit $\check{C}_1 \equiv \check{C}_2$

Extenzionalismus PL1

Varšava

hlavní město Polska

- | | | |
|---------------------|---|--|
| Varšava | – | jméno individua, jasně identifikovatelné a odlišitelné |
| hlavní město Polska | – | individuová role, momentálně identifikuje Varšavu, ale dříve to byl i Krakov |

'hlavní město Polska':

- závisí na světě a čase
- pochopení významu, ale není vázané na znalost obsahu – tj. význam na světě a čase *nezávisí*

číslo X je větší než číslo Y

budova X je větší než budova Y

- | | | |
|-----------------------|---|---|
| matematické větší než | – | relace dvojic čísel, pevně daná |
| empirické větší než | – | vztah dvou individuí, který se může měnit v čase (otec a syn) |

Extenzionalismus PL1

Varšava

hlavní město Polska

- Varšava – **jméno individua**, jasně identifikovatelné a odlišitelné
- hlavní město Polska – **individuová role**, momentálně identifikuje Varšavu, ale dříve to byl i Krakov

'*hlavní město Polska*':

- závisí na světě a čase
- pochopení významu, ale není vázané na znalost obsahu – tj. **význam** na světě a čase **nezávisí**

číslo X je větší než číslo Y

budova X je větší než budova Y

- matematické větší než – **relace** dvojic čísel, pevně daná
- empirické větší než – **vztah** dvou individuí, který se může měnit v čase (otec a syn)

Extenzionalismus PL1

Varšava

hlavní město Polska

- Varšava – **jméno individua**, jasně identifikovatelné a odlišitelné
- hlavní město Polska – **individuová role**, momentálně identifikuje Varšavu, ale dříve to byl i Krakov

'hlavní město Polska':

- závisí na světě a čase
- pochopení významu, ale není vázané na znalost obsahu – tj. **význam** na světě a čase **nezávisí**

číslo X je větší než číslo Y

budova X je větší než budova Y

matematické větší než – **relace** dvojic čísel, pevně daná

empirické větší než – **vztah** dvou individuí, který se může měnit v čase (otec a syn)

Extenzionalismus PL1

Varšava

hlavní město Polska

- Varšava – **jméno individua**, jasně identifikovatelné a odlišitelné
- hlavní město Polska – **individuová role**, momentálně identifikuje Varšavu, ale dříve to byl i Krakov

'hlavní město Polska':

- závisí na světě a čase
- pochopení významu, ale není vázané na znalost obsahu – tj. **význam** na světě a čase **nezávisí**

číslo X je větší než číslo Y

budova X je větší než budova Y

- matematické větší než – **relace** dvojic čísel, pevně daná
- empirické větší než – **vztah** dvou individuí, který se může měnit v čase (otec a syn)

Extenzionalismus PL1 – pokrač.

ano

V Brně prší

- ano* – pravdivostní hodnota *true*
- V Brně prší* – **propozice** – označuje pravdivostní hodnotu, která se mění (alespoň) v čase

i když hodnota někdy závisí na světě a čase, samotný význam na nich nezávisí

Extenze a intenze

Definujeme:

- **intenze** – objekty typu funkcí, jejichž hodnoty závisí na světě a čase
- **extenze** – ostatní objekty (na světě a čase nezávislé)

časté extenze a intenze:

<i>extenze</i>	<i>intenze</i>
individua	individuové role
třídy	vlastnosti
relace	vztahy
pravdivostní hodnoty	propozice
funkce	empirické funkce
čísla	veličiny

Extenze a intenze

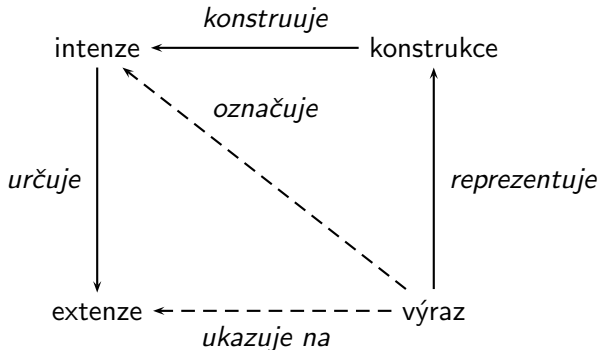
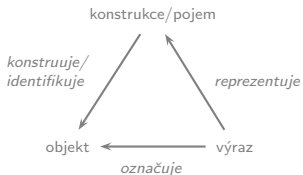
Definujeme:

- **intenze** – objekty typu funkcí, jejichž hodnoty závisí na světě a čase
- **extenze** – ostatní objekty (na světě a čase nezávislé)

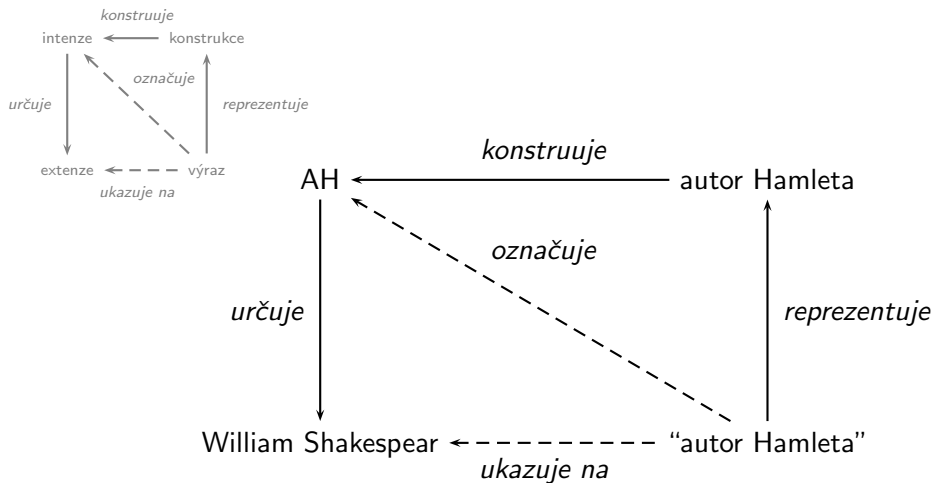
časté extenze a intenze:

<i>extenze</i>	<i>intenze</i>
individua	individuové role
třídy	vlastnosti
relace	vztahy
pravdivostní hodnoty	propozice
funkce	empirické funkce
čísla	veličiny

Rozšířený vztah výrazu a významu u intenzí



Rozšířený vztah výrazu a významu u intenzí



Transparentní intenzionální logika

- *Transparent Intensional Logic*, TIL
- **logický systém** speciálně navržený pro zachycení **významu výrazů PJ**
- autor **Pavel Tichý**: *The Foundations of Frege's Logic*, de Gruyter, Berlin, New York, 1988.
- obdobná teorie – *Montagueho intenzionální logika* – Tichý ukazuje její nedostatky
- Tichý vychází z myšlenek – *Gottlob Frege* (1848 – 1925, logik) a *Alonzo Church* (1903 – 1995, teorie typů)
- vlastnosti:
 - rozvětvená **typová hierarchie** (s typy **vyšších řádů**)
 - **temporální**
 - **intenzionální** (intenze \times extenze)
- **transparentost**:
 1. nositel významu (**konstrukce**) není prvek formálního aparátu, tento aparát pouze *studuje* konstrukce
 2. zachycení intenzionality je přesně popsáno z matematického hlediska

Obsah

- 1 Intenzionální sémantika
 - Logická analýza přirozeného jazyka
 - Nedostatečná expresivita PL1
 - Extenzionalismus PL1
 - Extenze a intenze
 - Transparentní intenzionální logika
- 2 Reprezentace znalostí
 - Reprezentace znalostí
 - Logika a sémantické sítě
 - Rámce
 - Pravidlové systémy
 - Nejistota a pravděpodobnost

otázka:

Jak zapíšeme znalosti o problému/doméně?

Když je zapíšeme, můžeme z nich mechanicky odvodit nová fakta?

- **reprezentace znalostí** (*knowledge representation*) – hledá způsob vyjádření znalostí počítačově zpracovatelnou formou (za účelem odvozování)
- **vyvozování znalostí** (*reasoning*) – zpracovává znalosti uložené v **bázi znalostí** (*knowledge base, KB*) a provádí **odvození** (*inference*) nových závěrů:
 - odpovědi na dotazy
 - zjištění faktů, které vyplývají z faktů a pravidel v KB
 - odvodit akci, která vyplývá z dodaných znalostí, ...

otázka:

Jak zapíšeme znalosti o problému/doméně?

Když je zapíšeme, můžeme z nich mechanicky odvodit nová fakta?

- **reprezentace znalostí** (*knowledge representation*) – hledá způsob vyjádření znalostí počítačově zpracovatelnou formou (za účelem odvozování)
- **vyvozování znalostí** (*reasoning*) – zpracovává znalosti uložené v **bázi znalostí** (*knowledge base, KB*) a provádí **odvození** (*inference*) nových závěrů:
 - odpovědi na dotazy
 - zjištění faktů, které vyplývají z faktů a pravidel v KB
 - odvodit akci, která vyplývá z dodaných znalostí, ...

otázka:

Jak zapíšeme znalosti o problému/doméně?

Když je zapíšeme, můžeme z nich mechanicky odvodit nová fakta?

- **reprezentace znalostí** (*knowledge representation*) – hledá způsob vyjádření znalostí počítačově zpracovatelnou formou (za účelem odvozování)
- **vyvozování znalostí** (*reasoning*) – zpracovává znalosti uložené v **bázi znalostí** (*knowledge base, KB*) a provádí **odvození** (*inference*) nových závěrů:
 - odpovědi na dotazy
 - zjištění faktů, které vyplývají z faktů a pravidel v KB
 - odvodit akci, která vyplývá z dodaných znalostí, ...

Reprezentace znalostí

proč je potřeba speciální **reprezentace znalostí**?

vnímání lidí × *vnímání počítačů*

- člověk

- když dostane novou věc (třeba pomeranč) – **prozkoumá** a **zapamatuje** si ho (a třeba sni)
- během tohoto procesu člověk zjistí a uloží všechny základní vlastnosti
- později, když se **zmíní** daná věc, vyhledají se a připomenou uložené informace

- počítač

- musí se spolehnout na informace od lidí
- jednodušší informace – přímé *programování*
- složité informace – zadané v **symbolickém jazyce**

Reprezentace znalostí

proč je potřeba speciální **reprezentace znalostí**?

vnímání lidí × *vnímání počítačů*

- člověk

- když dostane novou věc (třeba pomeranč) – **prozkoumá** a **zapamatuje** si ho (a třeba sni)
- během tohoto procesu člověk zjistí a uloží všechny základní vlastnosti
- později, když se **zmíní** daná věc, vyhledají se a připomenou uložené informace

- počítač

- musí se spolehnout na informace od lidí
- jednodušší informace – přímé *programování*
- složité informace – zadané v **symbolickém jazyce**

Reprezentace znalostí

proč je potřeba speciální **reprezentace znalostí**?

vnímání lidí × *vnímání počítačů*

- **člověk**

- když dostane novou věc (třeba pomeranč) – **prozkoumá** a **zapamatuje** si ho (a třeba sni)
- během tohoto procesu člověk zjistí a uloží všechny základní vlastnosti
- později, když se **zmíní** daná věc, vyhledají se a připomenou uložené informace

- **počítač**

- musí se spolehnout na informace od lidí
- jednodušší informace – přímé *programování*
- složité informace – zadané v **symbolickém jazyce**

Volba reprezentace znalostí

která **reprezentace znalostí** je **nejlepší**?

Pro řešení skutečně obtížných problémů musíme použít několik různých reprezentací. Důvodem pro to je to, že každý typ datových struktur má své přínosy i nedostatky a žádná z nich není adekvátní pro všechny různé funkce používané v tom, čemu říkáme "zdravý rozum" (common sense).

– Marvin Minsky

Volba reprezentace znalostí

která **reprezentace znalostí** je **nejlepší**?

Pro řešení skutečně obtížných problémů musíme použít několik různých reprezentací. Důvodem pro to je to, že každý typ datových struktur má své přínosy i nedostatky a žádná z nich není adekvátní pro všechny různé funkce používané v tom, čemu říkáme “zdravý rozum” (common sense).

– Marvin Minsky

Reprezentace znalostí pomocí logiky nebo sém. sítí

Logika:

- znalosti uloženy ve formě **logických formulí**
- vyvozování nových znalostí = hledání **důkazu**

Sémantické sítě:

- reprezentace faktových znalostí (pojmy + vztahy)
- znalosti jsou uloženy ve formě grafu
- nejdůležitější vztahy:
 - **podtřída** (*subclass*) – vztah mezi třídami
 - **instance** – vztah mezi konkrétním objektem a jeho rodičovskou třídou
- jiné vztahy – část (has-part), barva, ...

Reprezentace znalostí pomocí logiky nebo sém. sítí

Logika:

- znalosti uloženy ve formě **logických formulí**
- vyvozování nových znalostí = hledání **důkazu**

Sémantické sítě:

- reprezentace faktových znalostí (pojmy + vztahy)
- znalosti jsou uloženy ve formě grafu
- nejdůležitější vztahy:
 - **podtřída** (*subclass*) – vztah mezi třídami
 - **instance** – vztah mezi konkrétním objektem a jeho rodičovskou třídou
- jiné vztahy – část (has-part), barva, ...

Rámce

Rámce (*frames*):

- varianta sémantických sítí
- velice populární pro reprezentaci znalostí v expertních systémech
- všechny informace relevantní pro daný pojem se ukládají do univerzálních struktur – **rámců**
- stejně jako sémantické sítě, rámce podporují dědičnost
- OO programovací jazyky vycházejí z teorie rámců

Rámce – příklad

rámec obsahuje **objekty**, **sloty** a **hodnoty slotů**
 příklady rámců:

savec:

<i>podtřída:</i>	zvíře
<i>část:</i>	hlava
* <i>má_kožich:</i>	ano

slon:

<i>podtřída:</i>	savec
* <i>barva:</i>	šedá
* <i>velikost:</i>	velký

Nellie:

<i>instance:</i>	slon
<i>mít_rád:</i>	jablka

'*' označuje **vzorové hodnoty**, které mohou měnit hodnoty u podtříd a instancí

Sémantické sítě × rámce

sémantické sítě	rámce
uzly	objekty
spoje	sloty
uzel na druhém konci spoje	hodnota slotu

deskripční logika – logický systém, který manipuluje přímo s rámci

Pravidlové systémy

- snaha zachytit **produkčními pravidly** znalosti, které má expert
- obecná forma pravidel

IF *podmínka*
THEN *akce*

- podmínky – booleovské výrazy, dotazy na hodnoty **proměnných**
- akce – nastavení hodnot proměnných, příznaků, ...
- důležité vlastnosti:
 - znalosti mohou být strukturovány do modulů
 - systém může být snadno rozšířen přidáním nových pravidel beze změny zbytku systému

Pravidlové systémy

- snaha zachytit **produkčními pravidly** znalosti, které má expert
- obecná forma pravidel

IF *podmínka*
THEN *akce*

- podmínky – booleovské výrazy, dotazy na hodnoty **proměnných**
- akce – nastavení hodnot proměnných, příznaků, ...
- důležité vlastnosti:
 - znalosti mohou být strukturovány do modulů
 - systém může být snadno rozšířen přidáním nových pravidel beze změny zbytku systému

Pravidlová báze znalostí – příklad

pravidla pro **oblékání**:

- pravidlo 1 IF X je seriózní
AND X bydlí ve městě
THEN X by měl nosit sako
- pravidlo 2 IF X je akademik
AND X je společensky aktivní
AND X je seriózní
THEN X by měl nosit sako a
kravatu
- pravidlo 3 IF X bydlí ve městě
AND X je akademik
THEN X by měl nosit
kravatu
- pravidlo 4 IF X je podnikatel
AND X je společensky aktivní
AND X je seriózní
THEN X by měl nosit sako,
ale ne kravatu

společenská pravidla:

- pravidlo 5 IF X je podnikatel
AND X je ženatý
THEN X je společensky
aktivní
- pravidlo 6 IF X je akademik
AND X je ženatý
THEN X je seriózní

profesní pravidla:

- pravidlo 7 IF X učí na univerzitě
OR X učí na vysoké škole
THEN X je akademik
- pravidlo 8 IF X vlastní firmu
OR X je OSVČ
THEN X je podnikatel

Expertní systémy

- aplikace pravidlových systémů
- zaměřeny na specifické oblasti – medicínská diagnóza, návrh konfigurace počítače, expertíza pro těžbu ropy, ...
- snaha zachytit **znalosti experta** pomocí pravidel
ale znalosti experta zahrnují – postupy, strategie, odhady, ...
- expertní systém musí pracovat s procedurami, nejistými znalostmi, různými formami vstupu
- vhodné oblasti pro nasazení expertního systému:
 - **diagnóza** – hledání řešení podle symptomů
 - **návrh konfigurace** – složení prvků splňujících podmínky
 - **plánování** – posloupnost akcí splňujících podmínky
 - **monitorování** – porovnání chování s očekávaným chování, reakce na změny
 - **řízení** – ovládání složitého komplexu
 - **předpovědi** – projekce pravěpodobných závěrů z daných skutečností
 - **instruktáž** – inteligentní vyučování a zkoušení studentů

Nejistota a pravděpodobnost

definujme akci A_t jako “**Vyrazit na letiště t hodin před odletem letadla.**”
jak najít odpověď na otázku “*Dostanu se akcí A_t na letiště včas?*”

problémy:

1. částečná pozorovatelnost (stav vozovky, záměry ostatních řidičů, ...)
2. šum v senzorech (hlášení o dopravní situaci)
3. nejistota výsledků akcí (píchnutí kola, ...)
4. obrovská složitost modelování a předpovědi dopravní situace

čistě logický přístup tedy:

- riskuje chybu – “ A_5 mě tam dostane včas.”
- vede k závěrům, které jsou příliš slabé pro rozhodování: “ A_5 mě tam dostane včas, pokud nebude na dálnici nehoda a pokud nebude přšet a jestli nepíchnu kolo ...”

Nejistota a pravděpodobnost

definujme akci A_t jako “**Vyrazit na letiště t hodin před odletem letadla.**”
jak najít odpověď na otázku “*Dostanu se akcí A_t na letiště včas?*”

problémy:

1. částečná pozorovatelnost (stav vozovky, záměry ostatních řidičů, ...)
2. šum v senzorech (hlášení o dopravní situaci)
3. nejistota výsledků akcí (píchnutí kola, ...)
4. obrovská složitost modelování a předpovědi dopravní situace

čistě logický přístup tedy:

- riskuje chybu – “ A_5 mě tam dostane včas.”
- vede k závěrům, které jsou příliš slabé pro rozhodování: “ A_5 mě tam dostane včas, pokud nebude na dálnici nehoda a pokud nebude přšet a jestli nepíchnu kolo ...”

Nejistota a pravděpodobnost

definujme akci A_t jako “**Vyrazit na letiště t hodin před odletem letadla.**”
jak najít odpověď na otázku “*Dostanu se akcí A_t na letiště včas?*”

problémy:

1. částečná pozorovatelnost (stav vozovky, záměry ostatních řidičů, ...)
2. šum v senzorech (hlášení o dopravní situaci)
3. nejistota výsledků akcí (píchnutí kola, ...)
4. obrovská složitost modelování a předpovědi dopravní situace

čistě logický přístup tedy:

- riskuje chybu – “ A_5 mě tam dostane včas.”
- vede k závěrům, které jsou příliš slabé pro rozhodování: “ A_5 mě tam dostane včas, pokud nebude na dálnici nehoda a pokud nebude přšet a jestli nepíchnu kolo ...”

Metody pro práci s nejistotou

- defaultní/nemonotónní logika

Předpokládejme, že nepíchnu cestou kolo.

Předpokládejme, že A_5 bude OK, pokud se nenajde protipříklad.

- pravidla s faktory nejistoty

$A_5 \mapsto_{0.3}$ dostat se na letiště včas.

zalévání $\mapsto_{0.99}$ mokrý trávník

mokrý trávník $\mapsto_{0.7}$ déšť

- pravděpodobnost

Vzhledem k dostupným informacím, A_3 mě tam dostane včas s pravděpodobností 0.05.

Metody pro práci s nejistotou

- defaultní/nemonotónní logika

Předpokládejme, že nepíchnu cestou kolo.

Předpokládejme, že A_5 bude OK, pokud se nenajde protipříklad.

- pravidla s faktory nejistoty

$A_5 \mapsto_{0.3}$ dostat se na letiště včas.

zalévání $\mapsto_{0.99}$ mokrý trávník

mokrý trávník $\mapsto_{0.7}$ déšť

- pravděpodobnost

Vzhledem k dostupným informacím, A_3 mě tam dostane včas s pravděpodobností 0.05.

Metody pro práci s nejistotou

- defaultní/nemonotónní logika

Předpokládejme, že nepíchnu cestou kolo.

Předpokládejme, že A_5 bude OK, pokud se nenajde protipříklad.

- pravidla s faktory nejistoty

$A_5 \mapsto_{0.3}$ dostat se na letiště včas.

zalévání $\mapsto_{0.99}$ mokrý trávník

mokrý trávník $\mapsto_{0.7}$ déšť

- pravděpodobnost

Vzhledem k dostupným informacím, A_3 mě tam dostane včas s pravděpodobností 0.05.

Pravděpodobnost

běžná lidská tvrzení o pravděpodobnosti **shrnují** následky

- **lenosti** – nepodařilo se vypočítat všechny výjimky, podmínky, ...
- **neinformovanosti** – nedostatek relevantních údajů, počátečních podmínek, ...

subjektivní × **Bayesovská** pravděpodobnost:

- pravděpodobnostní vztah mezi tvrzením a jeho pravdivostí vzhledem k podmínkám:

$$P(A_5 | \text{žádné hlášené nehody}) = 0.5$$

nejedná se o vyjádření **pravděpodobnostní tendence** (ale může se získat ze znalostí podobných případů v minulosti)

- pravděpodobnost tvrzení se může měnit s novými (vstupními) podmínkami:

$$P(A_5 | \text{žádné hlášené nehody, je 4:00 ráno}) = 0.63$$

Pravděpodobnost

běžná lidská tvrzení o pravděpodobnosti **shrnují** následky

- **lenosti** – nepodařilo se vypočítat všechny výjimky, podmínky, ...
- **neinformovanosti** – nedostatek relevantních údajů, počátečních podmínek, ...

subjektivní × **Bayesovská** pravděpodobnost:

- pravděpodobnostní vztah mezi tvrzením a jeho pravdivostí vzhledem k podmínkám:

$$P(A_5 | \text{žádné hlášené nehody}) = 0.5$$

nejedná se o vyjádření **pravděpodobnostní tendence** (ale může se získat ze znalostí podobných případů v minulosti)

- pravděpodobnost tvrzení se může měnit s novými (vstupními) podmínkami:

$$P(A_5 | \text{žádné hlášené nehody, je 4:00 ráno}) = 0.63$$

Vyvozování z nejistých znalostí

- použití **náhodných proměnných** (*random variables*) – funkce, která vzorkům přiřazuje hodnoty
- **distribuce pravděpodobností** náhodné proměnné = (vektor) pravděpodobnost(i), že daná náhodná proměnná bude mít určitou konkrétní hodnotu

náh. prom. *Odd* – výsledek hodu kostkou bude lichý

náh. prom. *Weather* – jaké bude počasí (slunce, déšť, mraky, sníh)

$$Odd(1) = true \quad Weather(21.11.2005) = \text{déšť}$$

distribuce pravděpodobností proměnných *Odd* a *Weather*

$$P(Odd = true) = 1/6 + 1/6 + 1/6 = 1/2$$

$$P(Odd) = \langle 1/2, 1/2 \rangle$$

$$P(Weather) = \langle 0.72, 0.1, 0.08, 0.1 \rangle$$

- pravidla pro výpočet pravděpodobnosti logicky souvisejících událostí

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$

Bayesovské pravidlo pro vyvozování

pravidlo pro **podmíněnou pravděpodobnost** – $P(a|b) = \frac{P(a \wedge b)}{P(b)}$ if $P(b) \neq 0$
Bayesovské pravidlo pro určení **diagnostické** pravděpodobnosti ze znalosti **příčinné** pravděpodobnosti:

$$P(\text{Příčina}|\text{Následek}) = \frac{P(\text{Následek}|\text{Příčina})P(\text{Příčina})}{P(\text{Následek})}$$

např. ZMB zánět mozkových blan, ZK ztuhlý krk:

$$P(zmb|zk) = \frac{P(zk|zmb)P(zmb)}{P(zk)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008$$

vyvozování =

1. rozdělení akce na **atomické události**
2. zjištění pravděpodobností atomických událostí
3. výpočet/odvození pravděpodobností pomocí **složených distribucí pravděpodobností** (*joint probability distribution*)