Deskripční logika

Petr Křemen

FEL ČVUT

Co nás čeká

1. Základy deskripční logiky
2. Jazyk $\mathcal{ALC}$
   - Syntax a sémantika
3. Cyklické a acyklické TBOXy

Základy deskripční logiky

Kontrolní otázky k FOL[2]

- Co je to *term*, *axiom/formule*, *teorie*, *model*, *univerzální uzávěr*, *rezoluce*, *logický důsledek* ?
- Co vyjadřuje předpoklad otevřenosti (OWA)/uzavřenosti (CWA) světa ?
- Jaký je rozdíl, mezi predikátem (relací) a predikátovým symbolem ?
- Co znamená, řekneme-li, že FOL je *nerozhodnutelná* ?
- Co znamená, řekneme-li, že FOL je *monotónní* ?
- Co říká *věta o úplnosti*, *věta o korektnosti*,*věta o dedukci* ?

[2] First Order Logic = predikátová logika prvního řádu tak jak ji znáte.

Motivace

- Proč potřebujeme nový logický jazyk – cožpak nám nestačí FOL ?
  - FOL je nerozhodnutelná – mnoho logických důsledků nelze v konečném čase ověřit.
  - Leckdy nepotřebujeme plnou expresivitu FOL.
- Ale vždyť je tu přeci Prolog – používaná a optimalizovaná implementace FOL ?
  - Prolog není implementací FOL – OWA vs. CWA, místo logické negace používá NAF, problémy s vyjádřením disjunktivní znalosti, apod.
- Proč nám nestačí relační databáze ?
  - RDBMS používají předpoklad uzavřeného světa (CWA) a podporují pouze konečné domény.
  - RDBMS nejsou dostatečně flexibilní – změna v DB modelu bývá komplikovanější, než přidání/odebrání axiomu z ontologie.

Motivace (2)

- Proč si nevystačíme s rámci, sémantickými sítěmi, apod. ?
  - Nemají dostatečně dobře definovanou sémantiku:
  - Jaká je sémantika slotu - $\forall R \cdot C$ nebo $\exists R \cdot C$?
  - Jaká je sémantika definice tříd, $\sqsubseteq$ nebo $\equiv$?
- Co a k čemu jsou tedy deskripční logiky ?
  - jedná se o skupinu jazyků vytvořených pro modelování *terminologické znalosti*, *neúplné znalosti*. Jedná se téměř výhradně o podmnožiny FOL.
  - první jazyky vznikly jako snaha o formalizaci sémantických sítí a rámců. První implementace v 80's – systémy KL-ONE, KAON, Classic .

Koncepty a role

- Základními stavebními bloky deskripčních logik jsou :

  (atomické) koncepty - reprezentující (pojmenované) unární predikáty (třídy), např. *Rodic*, nebo $Osoba \sqcap \exists maDite \cdot Osoba$.

  (atomické) role - reprezentující (pojmenované) binární predikáty (relace), např. *maDite*, nebo $jeDitetem^{-}$

  individuály - reprezentující prvky tříd/jednotlivce, např. *JIRI*

- Ontologie $\mathcal{K}$ většiny deskripčních logik se skládá

  TBOXu $\mathcal{T}$ - reprezentující axiomy obecně platné v dané doméně, např. $\mathcal{T} = \{Muz \sqsubseteq Osoba\}$

  ABOXu $\mathcal{A}$ - reprezentující konkrétní relační strukturu, např. $\mathcal{A} = \{Muz(Jiri)\}$

- Jednotlivé deskripční logiky se liší v možnostech tvořit složitější koncepty, role a v typech axiomů.

Interpretace

…slouží k definování významu symbolů zavedených výše. Protože se pohybujeme v podmnožině FOL, definujeme interpretaci redukcí na FOL formule a využijeme standardní FOL strukturu.

Interpretace bude pro nás dvojice : $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, kde $\Delta^{\mathcal{I}}$ je interpretační doména a $\cdot^{\mathcal{I}}$ je interpretační funkce. Máme-li *atomický* koncept A, *atomickou* roli R a individuál a, potom

$$\begin{aligned} A^{\mathcal{I}} &\subseteq \Delta^{\mathcal{I}} \\ R^{\mathcal{I}} &\subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \\ a^{\mathcal{I}} &\in \Delta^{\mathcal{I}} \end{aligned}$$

Logický důsledek

Máme-li libovolnou množinu S axiomů (resp. ontologii $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, kde položíme $S = \mathcal{T} \cup \mathcal{A}$), potom

- $\mathcal{I} \models S$ pokud $\mathcal{I} \models \alpha$ pro všechny $\alpha \in S$ ($\mathcal{I}$ je modelem S, resp. $\mathcal{K}$)
- $S \models \beta$ pokud $\mathcal{I} \models \beta$ vždy když $\mathcal{I} \models S$ (β je logickým důsledkem S, resp. $\mathcal{K}$)
- S je konzistentní, pokud S má alespoň jeden model

Jazyk $\mathcal{ALC}$

$\mathcal{ALC}$ (= attributive language with complements)

Máme-li koncepty C, D, atomický koncept A a atomickou roli R, potom pro interpretaci $\mathcal{I}$:

koncept	*koncept*$^{\mathcal{I}}$	*popis*
$\top$	$\Delta^{\mathcal{I}}$	(univerzální koncept)
$\bot$	$\emptyset$	(nesplnitelný koncept)
$\neg C$	$\Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$	(negace)
$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$	(průnik)
$C \sqcup D$	$C^{\mathcal{I}} \cup D^{\mathcal{I}}$	(sjednocení)
$\forall R \cdot C$	$\{a \mid \forall b((a,b) \in R^{\mathcal{I}} \Rightarrow b \in C^{\mathcal{I}})\}$	(univ. omezení)
$\exists R \cdot C$	$\{a \mid \exists b((a,b) \in R^{\mathcal{I}} \wedge b \in C^{\mathcal{I}})\}$	(exist. omezení)

TBOX

axiom	$\mathcal{I} \models$ axiom iff	*popis*
$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$	(inkluze)
$C \equiv D$	$C^{\mathcal{I}} = D^{\mathcal{I}}$	(ekvivalence)

ABOX (UNA = unique name assumption[3])

axiom	$\mathcal{I} \models$ axiom iff	*popis*
$C(a)$	$a^{\mathcal{I}} \in C^{\mathcal{I}}$	(instance konceptu)
$R(a,b)$	$(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$	(instance role)

[3] dvě různé instance označují různé objekty domény

$\mathcal{ALC}$ – příklad

Příklad

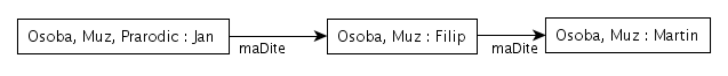
Uvažujme, že bychom chtěli vytvořit informační systém pro správu genealogických dat. Potřebujeme integrovat informace z různých zdrojů - databází, informačních systémů s *různými datovými modely*. Jako integrační vrstvu jsme si zvolili popis v deskripční logice. Máme atomické koncepty $Osoba$, Muz, $Prarodic$ a atomickou roli $maDite$.

- Jakým konceptem popíšeme množinu osob, které mají za své potomky pouze muže, nebo nemají potomky žádné ?
  - $Osoba \sqcap \forall maDite \cdot Muz$
- Jak definovat koncept $Prarodic$?
  - $Prarodic \equiv Osoba \sqcap \exists maDite \cdot \exists maDite \cdot \top$
- Jak by vypadal předchozí axiom v predikátové logice ?
  - $\forall x\,(Prarodic(x) \equiv (Osoba(x) \wedge \exists y\,(maDite(x,y) \wedge \exists z\,(maDite(y,z)))))$

Interpretace – Příklad

Příklad

- Uvažujme ontologii $\mathcal{K}_1 = (\{Prarodic \equiv Osoba \sqcap \exists maDite \cdot \exists maDite \cdot \top\}, \{Prarodic(JAN)\})$, modelem $\mathcal{K}_1$ může být např. interpretace $\mathcal{I}_1$:
    - $\Delta^{\mathcal{I}_1} = Muz^{\mathcal{I}_1} = Osoba^{\mathcal{I}_1} = \{Jan, Filip, Martin\}$
    - $maDite^{\mathcal{I}_1} = \{(Jan, Filip), (Filip, Martin)\}$
    - $Prarodic^{\mathcal{I}_1} = \{Jan\}$
    - $JAN^{\mathcal{I}_1} = \{Jan\}$
- tento model je konečný a má strukturu kořenového stromu s kořenem v uzlu Jan :



Vlastnosti modelů

Na konci předchozího příkladu jsme se dotkli velmi důležitých vlastností modelů deskripčních logik. Podívejme se na problém obecněji. Zvolený jazyk deskripční logiky má vlastnost

TMP (tree model property), pokud každý splnitelný koncept[4] C tohoto jazyka má model ve tvaru kořenového stromu.

FMP (finite model property), pokud každá konzistentní ontologie $\mathcal{K}$ tohoto jazyka má konečný model.

Obě tyto vlastnosti představují důležité charakteristiky příslušné deskripční logiky, které mají přímý vliv na inferenci (viz. dále).

Zejména (případně zobecněná) TMP je charakteristickou vlastností většiny deskripčních logik a je základem pro určení jejich rozhodnutelnosti a složitosti.

[4] Koncept je splnitelný, je-li alespoň jedním modelem interpretován jako neprázdná množina

Cyklické a acyklické TBOXy

Terminologické axiomy

definice je axiom ve tvaru $A \equiv C$, kde A je atomický koncept.

primitivní koncept se vyskytuje pouze na pravé straně terminologických axiomů

definovaný koncept je atomický koncept, který není primitivní

acyklický TBOX Mějme TBOX $\mathcal{T}$, který obsahuje pouze inkluze ($C \sqsubseteq D$). Utvořme graf, jehož vrcholy jsou všechny atomické koncepty vyskytující se v $\mathcal{T}$ a hrana vede od atomického konceptu A k B, pokud se A "vyskytuje" na leve straně nějakého axiomu a B na pravé straně téhož. $\mathcal{T}$ je acyklický, je-li tento graf acyklický.

Příklad

Příklad

primitivní (atomický) koncept

definovaný (atomický) koncept

$$\begin{aligned}
Woman &\equiv Person \sqcap Female \\
Man &\equiv Person \sqcap \neg Woman \\
Mother &\equiv Woman \sqcap \exists hasChild \cdot Person \\
Father &\equiv Man \sqcap \exists hasChild \cdot Person \\
Parent &\equiv Father \sqcup Mother \\
Grandmother &\equiv Mother \sqcap \exists hasChild \cdot Parent \\
MotherWithoutDaughter &\equiv Mother \sqcap \forall hasChild \cdot \neg Woman \\
Wife &\equiv Woman \sqcap \exists hasHusband \cdot Man
\end{aligned}$$

Příklad – pokračování

Každý acyklický TBOX $\mathcal{T}$ lze expandovat do sémanticky ekvivalentní $\mathcal{T}'$, která obsahuje pouze axiomy s pravou stranou tvořenou primitivním konceptem a levou stranou tvořenou pouze definovanými koncepty.

Příklad

primitivní koncept
definovaný koncept

$$\begin{aligned} Woman &\equiv Person \sqcap Female \\ Man &\equiv Person \sqcap \neg(Person \sqcap Female) \\ Mother &\equiv (Person \sqcap Female) \sqcap \exists hasChild \cdot Person \\ &\dots \end{aligned}$$

Z hlediska inference (viz. dále) je lze každý acyklický TBOX "zkompilovat" do ABOXu a není nutné pak TBOX vůbec uvažovat.

Příklad – CWA × OWA

Example

ABOX

$hasChild(JOCASTA, OEDIPUS)$ $\quad$ $hasChild(JOCASTA, POLYNEIKES)$
$hasChild(OEDIPUS, POLYNEIKES)$ $\quad$ $hasChild(POLYNEIKES, THERSANDROS)$
$Patricide(OEDIPUS)$ $\quad$ $\neg Patricide(THERSANDROS)$

Hrany reprezentují *hasChild* instance relací a barvy odlišují instance konceptů – *Patricide* a $\neg$*Patricide*

$$\begin{tikzcd}
JOCASTA \arrow[rr] \arrow[dr] & & POLYNEIKES \arrow[r] & THERSANDROS \\
& OEDIPUS \arrow[ur] & &
\end{tikzcd}$$

Dotaz 1 $(\exists hasChild \cdot (Patricide \sqcap \exists hasChild \cdot \neg Patricide))(JOCASTA)$,

Dotaz 2 Jaké jsou instance konceptu
$((\neg Patricide) \sqcap (\exists[^{-} hasChild]) \cdot Patricide \sqcap \exists hasChild^{-} \cdot \{JOCASTA\}$
? Jak by lišil výsledek, kdybychom uvažovali CWA ?

$$JOCASTA \longrightarrow x \longrightarrow y$$