Devátá přednáška

NAIL062 Výroková a predikátová logika

Jakub Bulín (KTIML MFF UK)

Zimní semestr 2024

Devátá přednáška

Program

- úvod do rezoluce v predikátové logice
- skolemizace
- grounding, Herbrandova věta

Materiály

Zápisky z přednášky, Sekce 8.1-8.3 z Kapitoly 8

Kapitola 8: Rezoluce v predikátové logice

8.1 Úvod

Rezoluce v predikátové logice

$T \models \varphi? \rightsquigarrow T \cup \{\neg\varphi\} \rightsquigarrow$ CNF formule $S \rightsquigarrow$ rezoluční zamítnutí

(pozor: φ musí být sentence!)

- literál je atomická formule $R(t_1, \ldots, t_n)$ nebo její negace
- klauzule je konečná množina literálů, formule množina klauzulí
- otevřenou formuli snadno převedeme do CNF, i univerzální kvantifikátor na začátku: $(\forall x)(P(x) \vee \neg Q(x)) \sim \{P(x), \neg Q(x)\}$
- co s existenčními kvantifikátory? nové symboly pro 'svědky'
  $(\exists x)(P(x) \vee \neg Q(x)) \rightsquigarrow \{P(c), \neg Q(c)\}$ "skolemizace"
- není ekvivalentní, ale zachovává [ne]splnitelnost, to nám stačí
- rezoluční krok? literály nemusí být stejné, stačí unifikovatelné
  z klauzulí $\{P(x), \neg Q(x)\}$ a $\{Q(f(c))\}$ odvodíme $\{P(f(c))\}$
- unifikace je substituce $\{x/f(c)\}$

Příklady

1. $T = \{(\forall x)P(x), (\forall x)(P(x) \to Q(x))\}$, $\varphi = (\exists x)Q(x)$

$$\neg\varphi = \neg(\exists x)Q(x) \sim (\forall x)\neg Q(x) \sim \neg Q(x)$$

$T \cup \{\neg\varphi\}$ je ekvivalentní $S = \{\{P(x)\}, \{\neg P(x), Q(x)\}, \{\neg Q(x)\}\}$
rezoluční zamítnutí: představte si p místo $P(x)$, q místo $Q(x)$

2. $T = \{(\forall x)(\exists y)R(x,y), R(x,y) \to Q(x)\}$, $\varphi = (\exists x)Q(x)$

$$T \cup \{\neg\varphi\} \sim \{(\forall x)(\exists y)R(x,y), \neg R(x,y) \vee Q(x), \neg Q(x)\}$$

formuli $(\forall x)(\exists y)R(x,y)$ nahradíme $R(x, f(x))$, kde f je nový unární funkční symbol (reprezentuje výběr svědka):

$$S = \{\{R(x, f(x))\}, \{\neg R(x,y), Q(x)\}, \{\neg Q(x)\}\}$$

není ekvivalentní, ale ekvisplnitelná (zde obě nesplnitelné), vidíme po substituci $y/f(x)$, která unifikuje $R(x, f(x))$ a $R(x,y)$

Unifikace

$$S = \{\{R(x, f(x))\}, \{\neg R(x, y), Q(x)\}, \{\neg Q(x)\}\}$$

- na úrovni výrokové logiky (ground level):
  $$\{\{r\}, \{\neg p, q\}, \{\neg q, p\}, \{\neg q\}\}$$
  není nesplnitelné! musíme využít, že $R(x, f(x))$ a $R(x, y)$ mají 'podobnou strukturu' (jsou unifikovatelné)
- klauzule $\{\neg R(x, y), Q(x)\}$ platí i po provedení libovolné substituce: $\{\neg R(x/t), Q(x/t)\}$ je důsledek S pro lib. term t
- představme si 'přidání' všech takto získaných klauzulí do S: potom už je na ground level nesplnitelné (ale nekonečné)
- unifikační algoritmus nám dá správnou substituci $y/f(x)$
- zahrneme už do rezolučního pravidla, tedy rezolventou klauzulí $\{P(c)\}$ a $\{\neg P(x), Q(x)\}$ bude klauzule $\{Q(c)\}$.

Rezoluční pravidlo

- zahrnuje aplikaci unifikace
- lze vybrat více literálů najednou, ale musí být unifikovatelné:

  např. z $\{R(x, f(x)), R(g(y), z)\}, \{\neg R(g(c), u), P(u)\}$ odvodíme rezolventu $\{P(f(g(c)))\}$ za použití unifikace

  $$\{x/g(c), y/c, z/f(g(c)), u/f(g(c))\}$$

- budeme vyžadovat disjunktní množiny proměnných v klauzulích; lze přejmenovat, proměnné mají lokální význam:

  $$\models (\forall x)(\psi \wedge \chi) \leftrightarrow (\forall x)\psi \wedge (\forall x)\chi$$

8.2 Skolemizace

Ekvisplnitelná otevřená teorie

- teorie T v jazyce L a T' v (ne nutně stejném) jazyce L' jsou ekvisplnitelné, pokud platí: T má model $\Leftrightarrow$ T' má model
- zajímá nás jen [ne]splnitelnost (dokazujeme sporem)
- pro převod do CNF a rezoluci potřebujeme otevřené formule

Cíl: Ke každé teorii T sestrojíme ekvisplnitelnou, otevřenou T'.

1. převod do prenexní normální formy (vytkneme kvantifikátory)
2. nahradíme generálními uzávěry (potřebujeme sentence!)
3. nahradíme sentence Skolemovými variantami (odstranění $\exists$)
4. odstraníme zbývající $\forall$, máme otevřené formule

Prenexní normální forma

Formule φ je v prenexní normální formě (PNF), je-li následujícího tvaru, kde $Q_i \in \{\forall, \exists\}$ a formule φ' je otevřená:

$$(Q_1 x_1) \ldots (Q_n x_n) \varphi'$$

- $(Q_1 x_1) \ldots (Q_n x_n)$ je kvantifikátorový prefix, φ' otevřené jádro
- univerzální formule: v PNF a všechny kvantifikátory jsou $\forall$

Tvrzení: Ke každé formuli φ existuje ekvivalentní formule v PNF.

Důkaz: nahrazujeme podformule ekvivalentními s cílem posunout kvantifikátory blíž kořeni Tree(φ), dle pravidel z násl. Lemmatu. $\square$

Důsledek: Existuje i ekvivalentní PNF sentence (generální uzávěr).

Pravidla vytýkání kvantifikátorů

Lemma: Označme $\overline{Q}$ opačný kvantifikátor ke Q. Jsou-li φ a ψ formule, kde x není volná v ψ, potom:

$$
\begin{aligned}
\neg(Qx)\varphi &\sim (\overline{Q}x)\neg\varphi \\
(Qx)\varphi \wedge \psi &\sim (Qx)(\varphi \wedge \psi) \\
(Qx)\varphi \vee \psi &\sim (Qx)(\varphi \vee \psi) \\
(Qx)\varphi \rightarrow \psi &\sim (\overline{Q}x)(\varphi \rightarrow \psi) \\
\psi \rightarrow (Qx)\varphi &\sim (Qx)(\psi \rightarrow \varphi)
\end{aligned}
$$

Důkaz: snadno ověříme sémanticky, nebo tablo metodou (potom ale nejsou-li sentence, musíme nahradit generálními uzávěry) $\square$

Pozorování: Nahradíme-li ve φ podformuli ψ ekvivalentní ψ', je i výsledná formule φ' ekvivalentní φ. (Připomeňme: $\varphi \sim \varphi'$ právě když mají stejné modely, tj. $\models \varphi \leftrightarrow \varphi'$)

Převod do PNF: příklad

$$(\forall z)P(x,z) \wedge P(y,z) \ \rightarrow \ \neg(\exists x)P(x,y)$$

$$\sim (\forall u)P(x,u) \wedge P(y,z) \ \rightarrow \ (\forall x)\neg P(x,y)$$

$$\sim (\forall u)(P(x,u) \wedge P(y,z)) \ \rightarrow \ (\forall v)\neg P(v,y)$$

$$\sim (\exists u)(P(x,u) \wedge P(y,z) \ \rightarrow \ (\forall v)\neg P(v,y))$$

$$\sim (\exists u)(\forall v)(P(x,u) \wedge P(y,z) \ \rightarrow \ \neg P(v,y))$$

- v prvním kroku přejmenujeme z na u, nesmí být volná v $P(y,z)$
- podobně ve druhém kroku x na v
- která pravidla používáme? sledujte postup na stromu formule
- chceme-li sentenci:

$$(\forall x)(\forall y)(\forall z)(\exists u)(\forall v)(P(x,u) \wedge P(y,z) \rightarrow \neg P(v,y))$$

Poznámky

1. proč se při vytýkání z antecedentu mění kvantifikátor?

$$\begin{aligned}(Qx)\varphi \to \psi &\sim \neg(Qx)\varphi \vee \psi \\ &\sim (\overline{Q}x)(\neg\varphi) \vee \psi \\ &\sim (\overline{Q}x)(\neg\varphi \vee \psi) \sim (\overline{Q}x)(\varphi \to \psi)\end{aligned}$$

2. proč nesmí být x volná v ψ? neplatilo by, např:

$$(\exists x)P(x) \wedge Q(x) \;\not\sim\; (\exists x)(P(x) \wedge Q(x))$$

musíme přejmenovat vázanou proměnnou x na novou:

$$(\exists x)P(x) \wedge Q(x) \;\sim\; (\exists y)P(y) \wedge Q(x) \sim (\exists y)(P(y) \wedge Q(x))$$

3. PNF není jednoznačná, lze vytýkat v různém pořadí; lepší je nejprve vytknout ty, ze kterých se nakonec stanou existenční:

$$(\exists y)(\forall x)\varphi(x,y) \text{ je lepší než } (\forall x)(\exists y)\varphi(x,y)$$

(protože “y nezávisí na x”)

Skolemova varianta

Je-li PNF sentence univerzální, tvaru $(\forall x_1)\dots(\forall x_n)\psi(x_1,\dots,x_n)$, nahradíme otevřeným jádrem ψ. Jinak musíme provést skolemizaci:

Buď φ L-sentence v PNF, všechny vázané proměnné různé. Nechť

- existenční kvantifikátory jsou $(\exists y_1),\dots,(\exists y_n)$ (v tom pořadí)
- pro každé i jsou $(\forall x_1),\dots,(\forall x_{n_i})$ právě všechny univerzální kvantifikátory předcházející $(\exists y_i)$ v prefixu φ

Buď L' rozšíření L o nové funkční symboly $f_1,\dots,f_n$, kde f_i je n_i-ární. Skolemova varianta φ je L'-sentence φ_S vzniklá odstraněním $(\exists y_i)$ a substitucí termu $f_i(x_1,\dots,x_{n_i})$ za y_i, postupně pro $i = 1,\dots,n$.

$$\varphi = (\exists y_1)(\forall x_1)(\forall x_2)(\exists y_2)(\forall x_3)\ R(y_1, x_1, x_2, y_2, x_3)$$
$$\varphi_S = (\forall x_1)(\forall x_2)(\forall x_3)\ R(f_1, x_1, x_2, f_2(x_1, x_2), x_3)$$

- musí být sentence! pro $(\exists y)E(x,y)$ ne $\cancel{E(x,c)}$ ale $E(x, f(x))$
- nové symboly! (jedinou rolí je reprezentovat 'svědky' ve φ)

Je to konzervativní extenze

Lemma: Buď φ L-sentence $(\forall x_1)\dots(\forall x_n)(\exists y)\psi$, f nový funkční symbol, a φ' sentence $(\forall x_1)\dots(\forall x_n)\psi(y/f(x_1,\dots,x_n))$. Potom:

(i) L-redukt každého modelu φ' je modelem φ, a

(ii) každý model φ lze expandovat na model φ'.

Důkaz: (i) Buď $\mathcal{A}'$ model φ', $\mathcal{A}$ jeho L-redukt, $e : \mathrm{Var} \to \mathcal{A}$.
$\mathcal{A} \models \varphi[e]$ platí neboť $\mathcal{A} \models \psi[e(y/a)]$ pro $a = (f(x_1,\dots,x_n))^{\mathcal{A}'}[e]$.

(ii) Protože $\mathcal{A} \models \varphi$, existuje funkce $f^A : A^n \to A$, že pro každé ohodnocení e platí $\mathcal{A} \models \psi[e(y/a)]$ pro $a = f^A(e(x_1),\dots,e(x_n))$.
To znamená, že expanze o funkci f^A splňuje φ'. $\square$

- říká, že $\{\varphi'\}$ je konzervativní extenze $\{\varphi\}$, opakovaná aplikace dává Skolemovu větu (výsledek skolemizace je otevřená konzervativní extenze, speciálně je ekvisplnitelná)
- expanze v (ii) není jednoznačná (na rozdíl od extenze o definici nového funkčního symbolu)

Skolemova věta (shrnutí postupu)

Věta: Každá teorie má otevřenou konzervativní extenzi.

Důkaz Mějme L-teorii T. Axiomy nahradíme generálními uzávěry a převedeme do PNF, máme ekvivalentní L-teorii T'. V ní každý axiom nahradíme jeho Skolemovou variantou.

Tím získáme teorii T'' v rozšířeném jazyce L'. Lemma říká:

- L-redukt každého modelu T'' je model T'
- každý model T' lze expandovat do L' na model T''

Neboli T'' je konzervativní extenzí T', tedy i T. Je axiomatizovaná univerzálními sentencemi, odstraníme kvantifikátorové prefixy (vezmeme jádra) a máme ekvivalentní otevřenou teorii T'''. $\square$

Důsledek: Ke každé teorii můžeme pomocí skolemizace najít ekvisplnitelnou otevřenou teorii. (A tu už snadno převedeme do CNF.)

8.3 Grounding

Grounding

- základní (ground) instance otevřené φ ve volných proměnných $x_1, \dots, x_n$ je $\varphi(x_1/t_1, \dots, x_n/t_n)$, kde vš. t_i jsou konstantní

  Herbrandova věta říká, že je-li otevřená teorie nesplnitelná, lze to doložit "na konkrétních prvcích": existuje konečně mnoho základních instancí axiomů, jejichž konjunkce je nesplnitelná

- např. pro $T = \{P(x, y) \vee R(x, y), \neg P(c, y), \neg R(x, f(x))\}$ substituujeme konstantní termy $\{x/c, y/f(c)\}$:

  $$(P(c, f(c)) \vee R(c, f(c))) \ \wedge \ \neg P(c, f(c)) \ \wedge \ \neg R(c, f(c))$$

- základní atomické sentence chápeme jako prvovýroky:

  $$(p_1 \vee p_2) \wedge \neg p_1 \wedge \neg p_2$$

- to už snadno zamítneme výrokovou rezolucí
- p_1 znamená "platí $P(c, f(c))$", p_2 znamená "platí $R(c, f(c))$"

Přímá redukce do výrokové logiky

Herbrandova věta + korektnost a úplnost výrokové rezoluce dává následující, neefektivní postup (S' je moc velká, i nekonečná):

1. $S \rightsquigarrow S' =$ množina všech základních instancí klauzulí z S
2. atomické sentence v S' chápeme jako prvovýroky
3. S nesplnitelná $\Leftrightarrow$ S' zamítnutelná 'na úrovni výrokové logiky'

Např. pro $S = \{\{P(x,y), R(x,y)\}, \{\neg P(c,y)\}, \{\neg R(x, f(x))\}\}$

$$S' = \{\{P(c,c), R(c,c)\}, \{P(c,f(c)), R(c,f(c))\}, \{P(f(c),c), R(f(c),c)\}, \dots,$$
$$\{\neg P(c,c)\}, \{\neg P(c,f(c))\}, \{\neg P(c,f(f(c)))\}, \{\neg P(c,f(f(f(c))))\}, \dots,$$
$$\{\neg R(c,f(c))\}, \{\neg R(f(c),f(f(c)))\}, \{\neg R(f(f(c)),f(f(f(c))))\}, \dots\}$$

S' je nesplnitelná obsahuje konečnou nesplnitelnou podmnožinu:

$$\{\{P(c,f(c)), R(c,f(c))\}, \{\neg P(c,f(c))\}, \{\neg R(c,f(c))\}\} \vdash_R \square$$

Efektivnější je hledat vhodné základní instance unifikací [za chvíli]

Herbrandův model

Mějme jazyk $L = \langle \mathcal{R}, \mathcal{F} \rangle$ s alespoň jedním konstantním symbolem.
L-struktura $\mathcal{A} = \langle A, \mathcal{R}^{\mathcal{A}}, \mathcal{F}^{\mathcal{A}} \rangle$ je Herbrandův model, jestliže:

- A je množina všech konst. L-termů (Herbrandovo univerzum)
- pro každý n-ární $f \in \mathcal{F}$ a (konstantní) "t_1", ..., "t_n" $\in A$:
  $$f^{\mathcal{A}}(\text{“}t_1\text{”}, \ldots, \text{“}t_n\text{”}) = \text{“}f(t_1, \ldots, t_n)\text{”}$$
- speciálně, pro konstantní symbol $c \in \mathcal{F}$ je $c^{\mathcal{A}} = \text{“}c\text{”}$
- na relační symboly neklademe podmínky

Např. $L = \langle P, f, c \rangle$ (P unární rel., f binární funkční, c konstantní)
Herbrandův model je každá struktura $\mathcal{A} = \langle A, P^{\mathcal{A}}, f^{\mathcal{A}}, c^{\mathcal{A}} \rangle$, kde

- $A = \{\text{“}c\text{”}, \text{“}f(c,c)\text{”}, \text{“}f(c, f(c,c))\text{”}, \text{“}f(f(c,c),c)\text{”} \ldots\}$
- $c^{\mathcal{A}} = \text{“}c\text{”}$
- $f^{\mathcal{A}}(\text{“}c\text{”}, \text{“}c\text{”}) = \text{“}f(c,c)\text{”}$, $f^{\mathcal{A}}(\text{“}c\text{”}, \text{“}f(c,c)\text{”}) = \text{“}f(c, f(c,c))\text{”}$, $f^{\mathcal{A}}(\text{“}f(c,c)\text{”}, \text{“}c\text{”}) = \text{“}f(f(c,c),c)\text{”}$, atd.
- $P^{\mathcal{A}} \subseteq A$ může být libovolná

Herbrandova věta

Věta (Herbrandova): Je-li T otevřená, v jazyce bez rovnosti a s alespoň jedním konstantním symbolem, potom:

- buď má T Herbrandův model, nebo
- existuje konečně mnoho základních instancí axiomů T, jejichž konjunkce je nesplnitelná.

Důkaz: $T_{\text{ground}} =$ množina všech základních instancí axiomů T

Zkonstruujeme "systematické tablo" τ z T_{ground} s $\mathrm{F}\bot$ v kořeni, ale z jazyka L, bez rozšíření o pomocné konstantní symboly na L_C. (Nepotřebujeme je, protože v T_{ground} nejsou kvantifikátory.)

Pokud má τ bezespornou větev, je "kanonický model" (opět bez pomocných symbolů) Herbrandovým modelem T.

Jinak je τ důkaz sporu, T_{ground} (a tedy i T) je nesplnitelná. Tablo τ je konečné, používá jen konečně mnoho $\alpha_{\text{ground}} \in T_{\text{ground}}$, jejich konjunkce už je nesplnitelná. $\square$

Poznámky

- konstatní symbol potřebujeme, aby existovaly vůbec nějaké konstantní termy (ale není-li v L žádný, můžeme ho přidat)
- Herbrandův model je podobný kanonickému, ale nepřidáváme pomocné symboly, a neříkáme nic o relacích
- je-li jazyk s rovností, najdeme Herbrandův model pro T^* (přidané axiomy rovnosti) a faktorizujeme podle $=^A$

Důsledky Herbrandovy věty

Důsledek: Je-li T otevřená v jazyce s konstantním symbolem, potom T má model, právě když má model teorie T_{ground}.

Důkaz: $\Rightarrow$ V modelu T platí i všechny základní instance axiomů. Je tedy i modelem T_{ground}.

$\Leftarrow$ Pokud T nemá model, podle Herbrandovy věty je nějaká konečná podmnožina teorie T_{ground} nesplnitelná. $\square$

Důsledek: Mějme otevřenou $\varphi(x_1, \ldots, x_n)$ v L s konst. symbolem. Potom existuje $m \in \mathbb{N}$ a konstantní L-termy t_{ij} $(i \in [m], j \in [n])$, že sentence $(\exists x_1) \ldots (\exists x_n)\varphi(x_1, \ldots, x_n)$ je pravdivá, právě když je následující formule (výroková) tautologie:

$$\varphi(x_1/t_{11}, \ldots, x_n/t_{1n}) \vee \cdots \vee \varphi(x_1/t_{m1}, \ldots, x_n/t_{mn})$$

Důkaz: Je pravdivá, právě když $(\forall x_1) \ldots (\forall x_n)\neg\varphi$ neboli $\neg\varphi$ je nesplnitelná. Stačí aplikovat Herbrandovu větu na $T = \{\neg\varphi\}$. $\square$