Lógica Computacional

Aula Teórica 17: Forma Normal de Skolem e Unificação

Ricardo Gonçalves

Departamento de Informática

10 de novembro de 2023

Forma Normal de Skolem

A fórmula está na FNCP e os quantificadores são todos universais.

Exemplos

- $Q(x) \vee P(x, y)$
- $\forall_x Q(f(x), y)$
- $\forall_x \forall_y P(x, f(y))$
- $\forall_x (P(g(x, y)) \wedge (Q(x) \vee P(x, f(x))))$

Contra-Exemplos

- $\exists_y P(x, y)$
- $\forall_x \exists_y f(x) = y$
- $\forall_x P(f(x), y) \vee (P(x, y) \wedge Q(y))$
- $\neg \forall_x (Q(x, y) \wedge P(x, f(x)))$

Forma Normal de Skolem

Definição

Uma fórmula φ da linguagem de primeira ordem está na Forma Normal de Skolem ou FNS, e escreve-se FNS(φ), se

$$\varphi = \forall_{x_1} \ldots \forall_{x_n} \psi$$

sendo ψ uma fórmula de primeira ordem sem quantificadores tal que FNC(ψ).

Quantificadores existenciais

E se a fórmula tiver quantificadores existenciais?

Podemos transformar em algo que não tenha existenciais?

Skolemização - Ideia intuitiva

Constantes de Skolem

Quando o quantificador existencial ocorre no início da fórmula: $\exists_x \forall_y P(x, y)$
Indica a existência de um elemento do domínio.
Qual? Não sabemos... mas existe.
Ideia: usar uma constante nova para representar esse elemento - constantes de Skolem.

Funções de Skolem

Quando o quantificador existencial tem universais antes, o elemento que existe depende dos valores quantificados universalmente.
Exemplo: $\forall_x \forall_y \exists_z \ Soma(x, y, z)$
Para todo o x e todo o y existe z (que depende da escolha de x e y)
Ideia: usar um símbolo de função binário novo que dado x e y devolve z correspondente - funções de Skolem

Forma Normal de Skolem

Procedimento de conversão

Seja δ tal que FNCP(δ). Vamos eliminar iterativamente cada um dos quantificadores existenciais de δ, da esquerda para a direita:

- Se δ é da forma

  $$\exists_x Q^1_{x_1} \ldots Q^n_{x_n} \psi$$

  então obtemos

  $$Q^1_{x_1} \ldots Q^n_{x_n} [\psi]^x_a$$

  onde a é uma constante nova (não ocorre em ψ)
- Para o quantificador existencial mais à esquerda:

  $$\forall_{x_1} \ldots \forall_{x_{i-1}} \exists_{x_i} Q^{i+1}_{x_{i+1}} \ldots Q^n_{x_n} \psi'$$

  obtemos

  $$\forall_{x_1} \ldots \forall_{x_{i-1}} Q^{i+1}_{x_{i+1}} \ldots Q^n_{x_n} [\psi]^{x_i}_{f(x_1,\ldots,x_{i-1})}$$

  onde f é um símbolo de função de aridade $i-1$ novo.

Forma Normal de Skolem

Dada φ uma fórmula de primeira ordem

1. Converter φ na FNCP (aula passada)
2. Eliminamos quantificadores existenciais (slide anterior)

O resultado, chamado de Skolemização de φ e representado por φ^S, está na FNS

Conversão para a Forma Normal de Skolem

Seja $\varphi = \neg(\forall_x \exists_y P(x,y,z) \vee \exists x \forall_y \neg Q(x,y,z))$

Como φ não está na FNCP, faz-se primeiro essa conversão.

$$
\begin{aligned}
\varphi &\equiv \neg\forall_x \exists_y P(x,y,z) \wedge \neg\exists_x \forall_y \neg Q(x,y,z) \text{ [Passo 2]} \\
&\equiv \exists_x \neg\exists_y P(x,y,z) \wedge \forall_x \neg\forall_y \neg Q(x,y,z) \text{ [Passo 2]} \\
&\equiv \exists_x \forall_y \neg P(x,y,z) \wedge \forall_x \exists_y \neg\neg Q(x,y,z) \text{ [Passo 2]} \\
&\equiv \exists_x \forall_y \neg P(x,y,z) \wedge \forall_x \exists_y Q(x,y,z) \text{ [Passo 2]} \\
&\equiv \exists_{x_1} \forall_{x_2} \neg P(x_1,x_2,z) \wedge \forall_{x_3} \exists_{x_4} Q(x_3,x_4,z) \text{ [Passo 3]} \\
&\equiv \exists_{x_1} \forall_{x_2} (\neg P(x_1,x_2,z) \wedge \forall_{x_3} \exists_{x_4} Q(x_3,x_4,z)) \text{ [Passo 4]} \\
&\equiv \exists_{x_1} \forall_{x_2} \forall_{x_3} \exists_{x_4} (\neg P(x_1,x_2,z) \wedge Q(x_3,x_4,z)) \text{ [Passo 4]}
\end{aligned}
$$

Está na FNCP. Vamos agora eliminar os quantificadores existenciais.

Conversão para a Forma Normal de Skolem

$$\varphi \equiv \exists_{x_1} \forall_{x_2} \forall_{x_3} \exists_{x_4} (\neg P(x_1, x_2, z) \wedge Q(x_3, x_4, z))$$

Eliminar quantificadores existencial - esquerda para a direita:

$$\exists_{x_1} \forall_{x_2} \forall_{x_3} \exists_{x_4} (\neg P(x_1, x_2, z) \wedge Q(x_3, x_4, z))$$

$\rightsquigarrow$

$$\forall_{x_2} \forall_{x_3} \exists_{x_4} (\neg P(a, x_2, z) \wedge Q(x_3, x_4, z))$$

e continuamos $\rightsquigarrow$

$$\forall_{x_2} \forall_{x_3} (\neg P(a, x_2, z) \wedge Q(x_3, f(x_2, x_3), z))$$

$\forall_{x_2} \forall_{x_3} (\neg P(a, x_2, z) \wedge Q(x_3, f(x_2, x_3), z))$ é a Skolemização de φ

Conversão para a Forma Normal de Skolem

Atenção: φ^S não é equivalente a φ

Basta comparar $\exists_x P(x)$ e $P(a)$:

Suponhamos que $M, \rho \Vdash \exists_x P(x)$
Então existe $u \in U$ tal que $M, \rho[x := u] \Vdash P(x)$
Mas... $\underline{a}_I$ pode não ser u
Logo, pode acontecer que $M, \rho \nVdash P(a)$

Apesar de não ser equivalente...

φ^S é possível se e só se φ é possível

φ^S é contraditória se e só se φ é contraditória

Resultado

Teorema de Skolem

Para qualquer fórmula de primeira ordem φ temos que:

1. FNS(φ^S)
2. φ é possível se e só se φ^S é possível

A prova faz-se por indução no número de quantificadores existenciais da fórmula.

Caso interessante - Teorema de Skolem

Seja $\varphi = \exists_x \psi$ com FNS(ψ). Então $\varphi^S = [\psi]^x_a$

Suponhamos que φ é possível.
Então existe estrutura de interpretação $\mathcal{M} = (U, I)$ e atribuição ρ tal que $\mathcal{M}, \rho \Vdash \exists_x \psi$.
Ou seja, existe $u \in U$ tal que $\mathcal{M}, \rho[x := u] \Vdash \psi$.
Podemos então considerar $\mathcal{M}' = (U, I')$ em tudo igual a $\mathcal{M}$, excepto, possivelmente, que $\underline{a}_{I'} = u$.
Logo $\mathcal{M}', \rho \Vdash [\psi]^x_a$, o que significa que $\varphi^S = [\psi]^x_a$ é possível.

Relembrando a resolução...

Exemplo em Primeira Ordem

Considere-se o seguinte conjunto de cláusulas, assumindo as variáveis universalmente quantificadas.

$$\{\{\neg Q(x,y), P(f(x),y)\}, \{\neg P(f(x),y), R(x,y,z)\}\}$$

- Um resolvente das duas cláusulas em cima é a cláusula $R_1 = \{\neg Q(x,y), R(x,y,z)\}$.
- Considere-se agora a cláusula $\{\neg P(z,y), R(x,y,z)\}$.
  Não conseguimos usar resolução directamente com a primeira cláusula do conjunto acima, mas substituindo $f(x)$ por z obtém-se a cláusula $\{\neg P(f(x),y), R(x,y,f(x))\}$
  para a qual já podemos encontrar um resolvente:
  $R_2 = \{\neg Q(x,y), R(x,y,f(x))\}$.
- Nota: R_2 é consequência de R_1. Se R_1 é satisfeita (para qualquer z), então é satisfeita para $z = f(x)$ (que é R_2).

Cláusulas de Primeira Ordem

Definição

Considere-se uma fórmula $\varphi \in F_{\Sigma}^{X}$ tal que FNS(φ), *i.e.*,

$$\varphi = \forall_{x_1} \ldots \forall_{x_n} \psi$$

sendo ψ uma fórmula sem quantificadores tal que FNC(ψ).

- Como todas as variáveis estão universalmente quantificadas (as variáveis livres estão *implicitamente* quantificadas), φ pode ser representada como um conjunto de cláusulas.
- Definimos $C(\varphi)$ como o conjunto das cláusulas que se obtêm de ψ (que está na FNC).

Cláusulas de Primeira Ordem

Resultados

- Para qualquer $\varphi \in F_{\Sigma}^{X}$ tal que FNS(φ), existe um único $C(\varphi)$ (a menos do nome das variáveis)

- Para quaisquer $\varphi, \psi \in F_{\Sigma}^{X}$, se $C(\varphi) = C(\psi)$ então $\varphi \equiv \psi$.

Estes resultados derivam dos respectivos da Lógica Proposicional.

Motivação

Considere:
$C_1 = \{\neg P(z, y), R(x, y, z)\}$ e $C_2 = \{\neg P(f(x), y), R(x, y, f(x))\}$.
No exemplo atrás, usámos uma substituição (z por $f(x)$) que converteu C_1 em C_2, o que permitiu encontrar resolvente com $C_3 = \{\neg Q(x, y), P(f(x), y)\}$.

Unificação - motivação

Para encontrar um resolvente de duas cláusulas é às vezes necessário encontrar substituições que tornem iguais duas fórmulas atómicas.

A esse processo se chama Unificação.

Substituição

Definição

Uma substituição é uma função $\sigma : X \to T_\Sigma^X$.

Dada $\sigma : X \to T_\Sigma^X$ uma substituição, $t \in T_\Sigma^X$ e $\varphi \in F_\Sigma^X$,

- $[t]^\sigma$ denota o termo que se obtém de t substituindo simultaneamente as suas variáveis de acordo com σ
- $[\varphi]^\sigma$ denota a fórmula que se obtém de φ substituindo simultaneamente as suas variáveis livres de acordo com σ.

Dadas duas substituições σ_1 e σ_2, por $\sigma_1\sigma_2$ representamos o resultado de aplicar σ_2 ao resultado de σ_1, isto é, para cada $x \in X$, temos que $\sigma_1\sigma_2(x) = [\sigma_1(x)]^{\sigma_2}$.

Substituição

Exemplo

Seja $\varphi = P(x, y) \rightarrow \exists_x Q(x)$ e
σ tal que $\sigma(x) = f(y)$ e $\sigma(y) = g(y)$, obtemos:

$$[\varphi]^\sigma = [P(x, y) \rightarrow \exists_x Q(x)]^{x,y}_{f(y),g(y)} = P(f(y), g(y)) \rightarrow \exists_x Q(x)$$

Atenção: substituir simultaneamente.

Quando queremos apenas indicar que valor uma substituição σ dá a algumas variáveis (o que na prática é o caso), usamos a notação:

$$\sigma = \{x_0/t_0, \ldots, x_n/t_n\}$$

e assumimos que $\sigma(x) = x$, para $x \notin \{x_0, \ldots, x_n\}$

Unificação

Definição

Um conjunto de literais $\mathcal{L}$ é unificável se existe uma substituição σ que aplicada a todos os elementos de $\mathcal{L}$ torna o conjunto singular (*i.e.*, os vários literais convertem-se num só). Nesse caso σ diz-se um unificador de $\mathcal{L}$.

Exemplo

Seja $\mathcal{L} = \{P(f(x), y), P(z, w)\}$.

- Como unificar?
- Tome-se $\sigma = \{z/f(x), y/w\}$. Obtemos:
- $[\{P(f(x), y), P(z, w)\}]^\sigma =$
  $\{[P(f(x), y)]^\sigma, [P(z, w)\}]^\sigma =$
  $\{P(f(x), w)\}$ – conjunto singular

Unificação

Unificações não são necessariamente únicas.

Considere-se novamente $\mathcal{L} = \{P(f(x), y), P(z, w)\}$.

- Já vimos que se escolhermos $\sigma = \{z/f(x), y/w\}$ obtemos:

$$[\mathcal{L}]^{\sigma} = \{P(f(x), w)\}$$

- Mas se escolhermos $\sigma' = \{z/f(x), w/y\}$ obtemos:

$$[\mathcal{L}]^{\sigma'} = \{P(f(x), y)\}$$

- E se escolhermos $\sigma'' = \{x/a,\ z/f(a),\ y/b,\ w/b\}$ obtemos:

$$[\mathcal{L}]^{\sigma''} = \{P(f(a), b)\}$$

Unificador mais geral

Definição

Dado um conjunto de literais $\mathcal{L}$, uma substituição σ^* é um unificador mais geral de $\mathcal{L}$, o que se denota por $umg(\mathcal{L})$, se

- σ^* é um unificador de $\mathcal{L}$
- qualquer outro unificador σ de $\mathcal{L}$ é tal que $\sigma^*\sigma = \sigma$.

Teorema

Um conjunto finito de literais é unificável se e só se tem um unificador mais geral.

Prova: Algoritmo de unificação.

Algoritmo de unificação

Dado um conjunto de literais $\mathcal{L}$

Ideia: Construir uma sequência $(\mathcal{L}_0, \sigma_0), \ldots, (\mathcal{L}_n, \sigma_n)$ tal que:

- $\mathcal{L}_n$ é singular
- $\sigma_0 \ldots \sigma_n$ é o unificador mais geral de $\mathcal{L}$

Se em algum passo não conseguimos construir $(\mathcal{L}_i, \sigma_i)$, podemos concluir que $\mathcal{L}$ não é unificável.

Algoritmo de unificação

Seja $\mathcal{L}$ um conjunto finito de literais.

- $(\mathcal{L}_0, \sigma_0) = (\mathcal{L}, \emptyset)$.

Para $k \geq 0$:

- Se $\mathcal{L}_k$ é singular então retornar:
  "$\sigma_0 \ldots \sigma_k$ é unificador mais geral de $\mathcal{L}$"
- Caso contrário:
  existem dois literais $L_i, L_j \in \mathcal{L}_k$ que diferem em pelo menos um símbolo. Seja n a menor posição em que L_i e L_j diferem.
  - Se a posição n num deles corresponde a uma variável z e no outro ao início de um termo t que não tem z, então:
    $\sigma_{k+1} = \{z/t\}$ e $\mathcal{L}_{k+1} = [\mathcal{L}_k]^{\sigma_{k+1}}$
  - Se algumas das hipótese acima não se verifica, então retornar:
    "$\mathcal{L}$ não é unificável"

Algoritmo de unificação: primeiro exemplo

Seja $\mathcal{L} = \{R(f(g(x)), a, x), R(f(g(b)), a, b), R(f(y), z, b)\}$.
$(\mathcal{L}_0, \sigma_0) = (\mathcal{L}, \emptyset)$

Tome-se $\sigma_1 = \{y/g(b)\}$ obtém-se

$$\mathcal{L}_1 = [\mathcal{L}_0]^{\sigma_1} = \{R(f(g(x)), a, x), R(f(g(b)), a, b), R(f(g(b)), z, b)\}$$

Como $\mathcal{L}_1$ não é singular, continuamos.
Tome-se $\sigma_2 = \{x/b\}$ e obtém-se

$$\mathcal{L}_2 = [\mathcal{L}_1]^{\sigma_2} = \{R(f(g(b)), a, b), R(f(g(b)), z, b)\}$$

Como $\mathcal{L}_2$ não é singular, continuamos.
Toma-se $\sigma_3 = \{z/a\}$ e obtém-se

$$\mathcal{L}_3 = [\mathcal{L}_2]^{\sigma_3} = \{R(f(g(b)), a, b)\}$$

Como $\mathcal{L}_3$ é singular, o unificador mais geral de $\mathcal{L}$ é $\sigma = \sigma_0\sigma_1\sigma_2\sigma_3$

Algoritmo de unificação: segundo exemplo

Seja $\mathcal{L} = \{R(f(g(x)), a, x), R(f(g(a)), a, b), R(f(y), a, b)\}$.

$(\mathcal{L}_0, \sigma_0) = (\mathcal{L}, \emptyset)$

Tomando $\sigma_1 = \{y/g(a)\}$ obtém-se

$$\mathcal{L}_1 = [\mathcal{L}_0]^{\sigma_1} = \{R(f(g(x)), a, x), R(f(g(a)), a, b)\}$$

Como $\mathcal{L}_1$ não é singular, procura-se nova substituição.

Tomando $\sigma_2 = \{x/a\}$ obtém-se

$$\mathcal{L}_2 = [\mathcal{L}_1]^{\sigma_2} = \{R(f(g(a)), a, a), R(f(g(a)), a, b)\}$$

Como $\mathcal{L}_2$ não é singular, procura-se nova substituição.

Como na posição onde os dois literais diferem nenhum deles tem uma variável, o algoritmo retorna "$\mathcal{L}$ não é unificável".

Algoritmo de unificação: terceiro exemplo

Seja $\mathcal{L} = \{R(f(g(x)), a, b), R(f(g(a)), a, b), R(f(x), a, b)\}$

$(\mathcal{L}_0, \sigma_0) = (\mathcal{L}, \emptyset)$

Tomando $\sigma_1 = \{x/g(a)\}$ obtém-se

$$\mathcal{L}_1 = [\mathcal{L}_0]^{\sigma_1} = \{R(f(g(g(a))), a, b), R(f(g(a)), a, b)\}$$

Como $\mathcal{L}_1$ não é singular, procura-se nova substituição.

Como na posição onde os dois literais diferem nenhum deles tem uma variável, o algoritmo retorna "$\mathcal{L}$ não é unificável".

Algoritmo de unificação: quarto exemplo

Seja $\mathcal{L} = \{P(x), P(f(x))\}$.

$(\mathcal{L}_0, \sigma_0) = (\mathcal{L}, \emptyset)$

Como $\mathcal{L}_0$ não é singular, procura-se substituição.

Na posição onde os dois literais diferem:

- um deles tem uma variável
- o outro tem o início de um termo que contém essa variável

O algoritmo retorna "$\mathcal{L}$ não é unificável".