C12

miercuri, 24 ianuarie 2024 23:24

Cum se face unificarea?

În logica predicatelor, rezoluția *unifică* un literal cu negatul lui:

$$A \vee P(t_1, t_2, \ldots, t_n) \quad \text{și} \quad B \vee \neg P(t_1', t_2', \ldots, t_n')$$

dacă putem unifica ("potrivi") argumentele lui P și $\neg P$: t_1 cu t_1',...

O *substituție* e o *funcție* care asociază unor *variabile* niște *termeni*:

$$\{x_1 \mapsto t_1, \ \ldots, x_n \mapsto t_n\}$$

Doi termeni se pot *unifica* dacă există o substituție care îi face egali

(o asemenea substituție se numește *unificator*)

$$f(x, g(y, z), t)\{x \mapsto h(z), y \mapsto h(b), t \mapsto u\} = f(h(z), g(h(b), z), u)$$

Termenul T cu substituția σ se notează uzual postfix: $T\sigma$

Substituția găsită se aplică predicatelor care rămân (rezolventul):

$$\frac{A \vee P(t_1, t_2, \ldots, t_n) \qquad B \vee \neg P(t_1', t_2', \ldots, t_n')}{A\sigma \vee B\sigma}$$

Reguli de unificare

O *variabilă* x poate fi unificată cu orice *termen* t (substituție) dacă x

O *variabilă* x poate fi unificată cu orice *termen* t (substituție) dacă x *nu apare* în t (altfel, substituind obținem un termen infinit)

deci nu: x cu $f(h(y), g(x, z))$; dar trivial, putem unifica x cu x

Doi *termeni* $f(\ldots)$ pot fi unificați doar dacă au aceeași funcție, și *argumentele* (termeni) pot fi unificate (poziție cu poziție)

Două *constante* (funcții cu 0 arg.) ⇒ unificate dacă sunt identice

Cât de generală e o demonstrație?

rezoluție:

$$\frac{\neg boy(x) \vee \neg girl(x) \vee child(x) \qquad boy(c)}{\neg girl(c) \vee child(c)}$$

Demonstrația e făcută *fără* a ține cont (sau înțelege) *semnificația* predicatelor *boy*, *child*, *good*, etc.:

puteau fi $P(x)$, $Q(x)$, $R(x)$, . . .

Demonstrația e valabilă pentru *orice predicate* care satisfac ipotezele.

Recapitulăm: sintaxa logicii predicatelor

Formulele logicii predicatelor sunt definite *structural recursiv*:

Termenii

variabilă v sau constantă c
$f(t_1, \cdot \cdot \cdot, t_n)$ cu f funcție n-ară și $t_1, \cdot \cdot \cdot, t_n$ *termeni*

Formule (well-formed formulașformule bine formate):

$P(t_1, \cdot \cdot \cdot, t_n)$	cu P predicat n-ar; $t_1, \cdot \cdot \cdot, t_n$ termeni
$\neg\alpha$	unde α este o formulă unde α,
$\alpha \rightarrow \beta$	β sunt formule
$\forall v\, \alpha$	cu v variabilă, α formulă: *cuantificare universală*
$t_1 = t_2$	cu t_1, t_2 termeni (în logica de ord. I cu egalitate)

În loc de propoziții avem *predicate* (peste *termeni*).

Sintaxă și semantică

Ca în logica propozițională (și pentru orice limbaj), deosebim

Ca în logica propozițională (și pentru orice limbaj), deosebim

sintaxa = *forma*, *regulile* după care construim ceva (aici, formule)

semantica = *înțelesul* construcțiilor de limbaj

La fel ca în logica propozițională, lucrăm cu

deducția (demonstrația): procedeu pur sintactic

implicația / consecința logică (consecința semantică):

interpretăm formula (*înțelesul* ei, valoarea de adevăr)

Ne interesează corespondența dintre aceste două aspecte.

Reguli și ce înseamnă aplicarea lor

Regulile discutate sunt *sintactice*:

manipulează forma (*simboluri*, nu înțelesul lor).

Regulile lui deMorgan: $\neg(a \vee b) = \neg a \wedge \neg b$, $\neg(a \wedge b) = \neg a \vee \neg b$

Înlocuim o formă cu alta.

Rezultatul: formulele sunt echivalente

Regulă: Dacă un literal L e singur într-o clauză:

ștergem toate clauzele din care apare

ștergem $\neg L$ din toate clauzele

ștergem toate clauzele din care apare

ștergem $\neg L$ din toate clauzele

Rezultatul: dacă formula era realizabilă, rămâne realizabilă

Axiomele calculului predicatelor

A1: $\alpha \to (\beta \to \alpha)$ (A1-A3 din logica propozițională)
A2: $(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$
A3: $(\neg\beta \to \neg\alpha) \to (\alpha \to \beta)$
A4: $\forall x\, (\alpha \to \beta) \to (\forall x\alpha \to \forall x\beta)$
A5: $\forall x\alpha \to \alpha[x \leftarrow t]$ dacă x poate fi substituit* cu t în α
A6: $\alpha \to \forall x\alpha$ dacă x nu apare liber în α

*Definim: putem *substitui* variabila x cu termenul t în $\forall y\phi$ dacă:
x nu apare liber în ϕ (substituția nu are efect) sau
x se poate substitui cu t în ϕ și y nu apare în t
(nu putem substitui variabile legate)

În logica cu egalitate, adăugăm si
A7: $x = x$
A8: $x = y \to \alpha = \beta$
unde β se obține din α înlocuind oricâte din aparițiile lui x cu y .

Regula de inferență: e suficient *modus ponens*:
$$\frac{A \qquad A \to B}{B}$$

Deducție

Fie H o mulțime de formule. O *deducție* (demonstrație) din H e un șir de formule $A_1, A_2, \cdots, A_n$, astfel ca $\forall i \in 1, n$

1. A_i este o *axiomă*, sau
2. A_i este o *ipoteză* (o formulă din H), sau
3. A_i rezultă prin *modus ponens* din A_j, A_k anterioare ($j, k < i$)

Spunem că A_n *rezultă* din H (e *deductibil*, e o *consecință*).

Notăm:

$$H \vdash A_n$$

Alte reguli de inferență

$$\frac{\forall x \ \phi(x)}{\phi(c)}$$ *instanțiere universală* (*vezi A5*)

unde c e o constantă arbitrară (nu apare anterior în demonstrație)
Dacă ϕ e valabil pentru orice x, atunci și pentru o valoare arbitrară c.

$$\frac{\phi(c)}{\forall x \ \phi(x)}$$ *generalizare universală* (*vezi A6*)

unde c e o valoare arbitrară (*nu apare în ipoteze*)
Dacă ϕ e valabilă pentru o valoare arbitrară, e valabilă pentru orice x.

unde *c* e o valoare arbitrară (*nu apare în ipoteze*)
Dacă ϕ e valabilă pentru o valoare arbitrară, e valabilă pentru orice x .

$$\frac{\exists x\ \phi(x)}{\phi(c)}$$ *instanțiere existențială*

Dacă există o valoare cu proprietatea ϕ, o instanțiem (cu un nume nou).

$$\frac{\phi(c)}{\exists x\ \phi(x)}$$ *generalizare existențială*

Dacă ϕ e adevărată pentru o valoare, există o valoare care o face adevărată

Definim noțiunile:

univers

interpretare

model

consecință semantică

Cum interpretăm o formulă ?

Cum interpretăm o formulă ?

Intuitiv, găsim un *înțeles* pentru fiecare simbol din formulă:

O *interpretare* (*structură*) I în logica predicatelor constă din:

- o mulțime nevidă U numită *universul* sau *domeniul* lui I (mulțimea valorilor pe care le pot lua variabilele)
- pentru orice simbol de constantă c, o valoare $c_I \in U$
- pentru orice simbol de funcție n-ară f, o funcție $f_I : U^n \to U$
- pentru orice simbol de predicat n-ar P, o submulțime $P_I \subseteq U^n$. (o *relație* n-ară pe U)

Deci, dăm o *interpretare* fiecărui simbol din formulă.

O interpretare *nu* dă valori variabilelor (vezi ulterior: atribuire).

Exemple de interpretări

$\forall x\, \forall y\, \forall z . P(x, y) \wedge P(y, z) \to P(x, z)$ tranzitivitate

De exemplu:

universul U = numere reale;

predicatul P: relația $\leq$

universul U = numere reale,
predicatul P: relația $\leq$

$\exists e \forall x \neg A(x, e)$

existența mulțimii vide:
predicatul $A(x, y)$ e $x \in y$

Implicația logică (consecința semantică)

Fie H o mulțime de formule și C o formulă.
Notăm $I \models H$ dacă I e un model pentru fiecare formulă din H.

Spunem că H *implică* C ($H \models C$) dacă pentru orice interpretare I,

$$I \models H \text{ implică } I \models C$$

(C e adev. în orice interpretare care satisface toate ipotezele din H)

Consistență și completitudine

La fel ca în logica propozițională
deducția (demonstrația) se face pur sintactic
consecința/ implicația logică e o noțiune semantică,
considerând *interpretări* și valori de adevăr.

Calculul predicatelor de ordinul I este *consistent* și *complet*:

$$H \vdash C \text{ dacă și numai dacă } H \models C$$

Concluzia C se poate *deduce* (demonstra) $\vdash$ din ipotezele H dacă și numai dacă ea e o *consecință semantică* $\models$ a ipotezelor H (e adevărată în orice *interpretare* care satisface toate ipotezele)

Dar: relația de implicație logică e doar *semidecidabilă*
dacă o formulă e o tautologie, ea poate fi demonstrată
dar dacă nu e, încercarea de a o demonstra (sau o refuta) poate continua la nesfârșit

Există logici mai bogate decât logica predicatelor

Principiul *inducției* matematice e (în ciuda numelui)
o *regulă de deducție* în *teoria aritmetică* a numerelor naturale

o *regulă de deducție* în *teoria aritmetică* a numerelor naturale

$$\forall P[P(0) \wedge \forall k \in \mathrm{N}.P(k) \rightarrow P(k+1)] \rightarrow \forall n \in \mathrm{N}\, P(n)$$

formulă în logica de *ordinul 2* (cuantificare peste predicate)

Logica are limitări

Teoria numerelor naturale cu adunare (aritmetica Presburger) e *decidabilă* (orice putem *exprima* despre adunarea numerelor naturale e *demonstrabil*).

Dar: nu putem exprima divizibilitate, numere prime, etc.

Aritmetica lui Peano (cu adunare și înmulțire) e mai bogată dar e *nedecidabilă*: sunt afirmații despre care *nu se poate decide* dacă sunt adevărate sau nu.

Teoremele de incompletitudine ale lui Gödel

Teoremele de incompletitudine ale lui Gödel

Prima teoremă de incompletitudine:

Orice sistem logic *consistent* care poate exprima aritmetica elementară e *incomplet*

i.e., se pot scrie afirmații care nu pot fi nici demonstrate nici infirmate în acel sistem

Demonstrație: codificând formule și demonstrații ca numere construim un număr care exprimă că formula sa e nedemonstrabilă

A doua teoremă de incompletitudine:

Consistența unui sistem logic capabil să exprime aritmetica elementară *nu poate fi demonstrată* în cadrul acelui sistem.

dar ar putea fi eventual demonstrată în alt sistem logic

Exercițiu

Formalizați în logica predicatelor:

1. Toti cainii latra noaptea.
2. Oricine are o pisica nu o sa aibe nici un

1. Toti cainii latra noaptea.
2. Oricine are o pisica nu o sa aibe nici un soarece.
3. Cei care adorm greu nu au nimic care latra noaptea.
4. Ionut are fie o pisica, fie un caine.
5. (Concluzia) Daca Ionut adoarme greu, atunci Ionut nu are nici un soarece.