

# Logică și structuri discrete

## Logica predicatelor

Casandra Holotescu

casandra@cs.upt.ro

<https://tinyurl.com/lecturesLSD>

În cursul de azi

Unificarea

Ce e o “demonstrație” și “totdeauna adevărat”  
și limitele a ce putem demonstra

Unificarea

## Cum se face unificarea?

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

$$A \vee P(t_1, t_2, \dots, t_n) \quad \text{și} \quad B \vee \neg P(t'_1, t'_2, \dots, t'_n)$$

dacă putem unifica (“potrivi”) argumentele lui  $P$  și  $\neg P$ :  $t_1$  cu  $t'_1, \dots$

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

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

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  $\sigma$  se notează uzual postfix:  $T\sigma$

Substituția găsită se aplică predicatelor care rămân (rezolventul):

$$\frac{A \vee P(t_1, t_2, \dots, t_n) \quad B \vee \neg P(t'_1, t'_2, \dots, t'_n)}{A\sigma \vee B\sigma}$$

## Reguli de unificare

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

Două *constante* (funcții cu 0 arg.)  $\Rightarrow$  unificate dacă sunt identice

$\Rightarrow$  cu aceste reguli, putem găsi *cel mai general unificator*  
(orice alt unificator se poate obține din el printr-o altă substituție)

## Demonstrații în logica predicatelor

## Cât de generală e o demonstrație?

$$\forall x(\text{boy}(x) \vee \text{girl}(x) \rightarrow \text{child}(x))$$

$$\neg \text{boy}(x) \vee \neg \text{girl}(x) \vee \text{child}(x)$$

$$\neg(\forall x \neg(\text{child}(x) \wedge \text{getstrain}(x))) \rightarrow \forall x \neg(\text{boy}(x) \wedge \text{good}(x))$$

$$(\neg \text{child}(y) \vee \neg \text{getstrain}(y)) \wedge \text{boy}(c) \wedge \text{good}(c)$$

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

*rezoluție:*

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.

# La ce e bună o demonstrație generală?

Exemplu de teoremă:

O *relație de echivalență* definește o *partiție* a mulțimii de definiție.

$\forall x R(x, x)$	reflexivitate
$\wedge \forall x \forall y (R(x, y) \rightarrow R(y, x))$	simetrie
$\wedge \forall x \forall y \forall z (R(x, y) \wedge R(y, z) \rightarrow R(x, z))$	tranzitivitate
$\rightarrow \forall x \forall y (\neg \exists z (R(x, z) \wedge R(y, z))$	clase disjuncte
$\vee \forall z (R(x, z) \leftrightarrow R(y, z)))$	sau clase identice

Teorema e adevărată indiferent de *semnificația* atribuită lui  $R(x, y)$  (indiferent de *interpretare*).  $R$  ar putea fi (printre altele):

$$R(x, y) \stackrel{\text{def}}{=} x \equiv y \pmod{d} \quad (\text{același rest la împărțirea cu } d)$$

$$R(x, y) \stackrel{\text{def}}{=} \text{length}(x) = \text{length}(y) \quad \text{liste/șiruri de aceeași lung.}$$

Definim mai precis ce înseamnă o *demonstrație*.



# Recapitulăm: sintaxa logicii predicatelor

Formulele logicii predicatelor sunt definite *structural recursiv*:

## Termenii

variabilă  $v$  sau constantă  $c$

$f(t_1, \dots, t_n)$  cu  $f$  funcție  $n$ -ară și  $t_1, \dots, t_n$  termenii

**Formule** (well-formed formulas, formule bine formate):

$P(t_1, \dots, t_n)$  cu  $P$  predicat  $n$ -ar;  $t_1, \dots, t_n$  termenii

$\neg \alpha$  unde  $\alpha$  este o formulă

$\alpha \rightarrow \beta$  unde  $\alpha, \beta$  sunt formule

$\forall v \alpha$  cu  $v$  variabilă,  $\alpha$  formulă: *cuantificare universală*

$t_1 = t_2$  cu  $t_1, t_2$  termenii (în logica de ord. I cu egalitate)

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

## Sintaxă și semantică

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 logic 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

*Rezultatul*: dacă formula era realizabilă, rămâne realizabilă

*Rezoluția*: 
$$\frac{p \vee A \quad \neg p \vee B}{A \vee B}$$

*Rezultatul*: dacă formula era realizabilă, rămâne realizabilă

## Axiomele calculului predicatelor

A1:  $\alpha \rightarrow (\beta \rightarrow \alpha)$  (A1-A3 din logica propozițională)

A2:  $(\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow ((\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \gamma))$

A3:  $(\neg\beta \rightarrow \neg\alpha) \rightarrow (\alpha \rightarrow \beta)$

A4:  $\forall x(\alpha \rightarrow \beta) \rightarrow (\forall x\alpha \rightarrow \forall x\beta)$

A5:  $\forall x\alpha \rightarrow \alpha[x \leftarrow t]$  dacă  $x$  poate fi substituit\* cu  $t$  în  $\alpha$

A6:  $\alpha \rightarrow \forall x\alpha$  dacă  $x$  nu apare liber în  $\alpha$

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

În logica cu egalitate, adăugăm și A7:  $x = x$

A8:  $x = y \rightarrow \alpha = \beta$

unde  $\beta$  se obține din  $\alpha$  înlocuind oricâte din aparițiile lui  $x$  cu  $y$ .

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

# Deducție

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

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

$$\frac{\forall x \varphi(x)}{\varphi(c)} \quad \textit{instanțiere universală} \textit{ (vezi A5)}$$

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

$$\frac{\varphi(c)}{\forall x \varphi(x)} \quad \textit{generalizare universală} \textit{ (vezi A6)}$$

unde  $c$  e o valoare arbitrară (nu apare în ipoteze)  
Dacă  $\varphi$  e valabilă pentru o valoare arbitrară, e valabilă pentru orice  $x$ .

$$\frac{\exists x \varphi(x)}{\varphi(c)} \quad \textit{instanțiere existențială}$$

Dacă există o valoare cu proprietatea  $\varphi$ , o instanțiem (cu un nume nou).

$$\frac{\varphi(c)}{\exists x \varphi(x)} \quad \textit{generalizare existențială}$$

Dacă  $\varphi$  e adevărată pentru o valoare, există o valoare care o face adevărată

# Semantica

Definim noțiunile:

univers

interpretare

model

consecință semantică

## 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  $f_I : U^n \rightarrow U$
  - 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

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

tranzitivitate

De exemplu:

universul  $U$  = numere reale;

predicatul  $P$ : relația  $\leq$

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

existența mulțimii vide:

predicatul  $A(x, y)$  e  $x \in y$

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

Putem găsi o interpretare în care e adevărată și una în care e falsă

## Atribuiri și valori de adevăr

Fie  $I$  o *interpretare* cu univers  $U$   
și fie  $V$  mulțimea tuturor simbolurilor de variabile.

O *atribuire* este o funcție  $s : V \rightarrow U$   
(dă fiecărei *variabile libere* o *valoare* din *univers*)

$\Rightarrow$  din atribuirea  $s$  se poate obține valoarea pentru orice *termen*  
(știm valoarea fiecărei variabile și înțelesul fiecărei funcții)

Interpretarea  $I$  dă și înțelesul fiecărui *predicat*

$\Rightarrow$  putem calcula *valoarea de adevăr* a unei formule  
 $\neg$ ,  $\rightarrow$  etc. au înțelesul cunoscut din logica propozițională  
trebuie definit înțelesul (semantica) lui  $\forall$

Spunem că  $\forall x \varphi$  e adevărată în interpretarea  $I$  cu atribuirea  $s$  dacă  
 $\varphi$  e adevărată înlocuind  $x$  cu *orice* valoare  $d \in U$  din univers.

## Modele și tautologii

Un *model* pentru o formulă  $\varphi$  e o interpretare în care formula e adevărată *pentru orice atribuire* a variabilelor.

Spunem că  $\varphi$  e *adevărată* în interpretarea  $I$ , și notăm  $I \models \varphi$

Obs: Dacă o formulă nu are variabile *libere*, valoarea ei de adevăr depinde doar de interpretare, nu și de vreo atribuire.

Def: O *tautologie* e o formulă adevărată în *orice* interpretare.

Spre deosebire de logica propozițională, în logica predicatelor, numărul interpretărilor e *infini*

⇒ nu mai putem construi exhaustiv tabelul de adevăr.

E *esențial* deci să putem *demonstra* o formulă (pornind de la axiome și reguli de inferență)

## 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

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

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

Poate fi definit ca o *schemă de axiome* (o axiomă pentru fiecare predicat), fără a cuantifica peste predicate

$$\forall \bar{x}[P(0, \bar{x}) \wedge \forall n(P(n, \bar{x}) \rightarrow P(S(n), \bar{x})) \rightarrow \forall nP(n, \bar{x})]$$

$\bar{x}$ : toate celelalte variabile de care depinde predicatul  $P$

Mai general: *inducția structurală*: demonstrăm proprietăți despre obiecte tot mai complexe (pt. obiecte definite inductiv/recursiv)

## 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

*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 aceluși sistem.

dar ar putea fi eventual demonstrată în alt sistem logic