Fundamentele limbajelor de programare

C08

Denisa Diaconescu
Traian Șerbănuță

Departamentul de Informatică, FMI, UB

Lambda calcul cu tipuri simple și unificare (recap)

Ce problemă am rezolvat în cursul trecut?

Type Inference

Pentru un lambda termen M fără tipuri, am adnotat termenul M cu tipuri obținând $\overline{M}$ și am rezolvat (parțial) problema

$$? \vdash \overline{M} : ?$$

(am găsit un context și un tip, pentru a avea o judecată legală).

Type Inference

Fie M un lambda termen fără tipuri.

Construim un context Γ_M pentru M:

$$\Gamma_M = \{x : X \mid x \in FV(M)\}$$

(toate variabilele de tip X introduse mai sus sunt noi și distincte)

Adnotăm M cu tipuri pentru variabilele legate obținând $\overline{M}$ prin inducție după structura lui M astfel:

- dacă $M = x$, atunci $\overline{M} = M$
- dacă $M = M_1\, N_1$, atunci $\overline{M} = \overline{M_1}\, \overline{N_1}$
- dacă $M = \lambda x. N$, atunci $\overline{M} = \lambda x : X. \overline{N}$, unde X este o variabilă de tip nouă

Sistemul $\lambda\!\rightarrow$ cu constrângeri

$$\Gamma \vdash M : \sigma \triangleright C$$

$$\frac{}{\Gamma \cup \{x : \tau\} \vdash x : \sigma \triangleright \{\sigma \doteq \tau\}}\ (var^*)$$

$$\frac{\begin{array}{c}\Gamma, x : \sigma \quad \vdash \quad M : \tau' \ \triangleright \ C' \\ C = C' \cup \{\tau \doteq \sigma \rightarrow \tau'\}\end{array}}{\Gamma \vdash (\lambda x : \sigma . M) : \tau \triangleright C}\ (\rightarrow_I^*)$$

$$\frac{\begin{array}{c}\Gamma \vdash M : \tau_1 \triangleright C_1 \qquad \Gamma \vdash N : \tau_2 \triangleright C_2 \\ C = C_1 \cup C_2 \cup \{\tau_1 \doteq \tau_2 \rightarrow \tau\}\end{array}}{\Gamma \vdash M\, N : \tau \triangleright C}\ (\rightarrow_E^*)$$

$\sigma, \tau, \tau', \tau_1, \tau_2$ variabile de tip

Sistemul $\lambda\to$ cu constrângeri

O judecată de forma $\Gamma \vdash M : \sigma \triangleright C$ este legală dacă constrângerile din C au o "soluție".

Fie M un lambda termen fără tipuri. Dacă există o constrângere de tipuri C_M și o variabilă de tip nouă V astfel încât

$$\Gamma_M \vdash \overline{M} : V \triangleright C_M$$

este o judecată legală, atunci M este *typable*.
(soluția o găsim prin C_M)

Type Inference - Exemplul 1

Fie $M_1 = (\lambda z.\,\lambda u.\,z)\,(y\,x)$.
Obținem $\Gamma_{M_1} = \{x : X, y : Y\}$ și $\overline{M_1} = (\lambda z : Z.\,\lambda u : U.\,z)\,(y\,x)$.

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\Gamma_{M_1} \cup \{z : Z, u : U\} \vdash z : \delta \triangleright D \\ C_1' = D \cup \{\tau_1' \doteq (U \to \delta)\}\end{array}}{\Gamma_{M_1} \cup \{z : Z\} \vdash \lambda u : U.\,z : \tau_1' \triangleright C_1'}\,(\to_I^*) \\ C_1 = C_1' \cup \{\tau_1 \doteq (Z \to \tau_1')\}}{\Gamma_{M_1} \vdash \lambda z : Z.\,\lambda u : U.\,z : \tau_1 \triangleright C_1}\,(\to_I^*) \qquad \dfrac{\begin{array}{c}\Gamma_{M_1} \vdash y : \sigma_1 \triangleright C_2' \quad \Gamma_{M_1} \vdash x : \sigma_2 \triangleright C_2'' \\ C_2 = C_2' \cup C_2'' \cup \{\sigma_1 \doteq \sigma_2 \to \tau_2\}\end{array}}{\Gamma_{M_1} \vdash y\,x : \tau_2 \triangleright C_2}\,(\to_E^*) \quad C_{M_1} = C_1 \cup C_2 \cup \{\tau_1 \doteq (\tau_2 \to V)\}}{\Gamma_{M_1} \vdash (\lambda z : Z.\,\lambda u : U.\,z)\,(y\,x) : V \triangleright C_{M_1}}\,($$

$D = \{\delta \doteq Z\}$
$C_1' = \{\delta \doteq Z, \tau_1' \doteq (U \to \delta)\}$
$C_1 = \{\delta \doteq Z, \tau_1' \doteq (U \to \delta), \tau_1 \doteq (Z \to \tau_1')\}$

$C_2' = \{\sigma_1 \doteq Y\}$
$C_2'' = \{\sigma_2 \doteq X\}$
$C_2 = \{\sigma_1 \doteq Y, \sigma_2 \doteq X, \sigma_1 \doteq \sigma_2 \to \tau_2\}$

$$C_{M_1} = \{\delta \doteq Z, \tau_1' \doteq (U \to \delta), \tau_1 \doteq (Z \to \tau_1'), \sigma_1 \doteq Y, \sigma_2 \doteq X,$$
$$\sigma_1 \doteq \sigma_2 \to \tau_2, \tau_1 \doteq (\tau_2 \to V)\}$$

Constrângerile C_{M_1} au "soluție". Ce înseamnă asta?

Type Inference - Exemplul 2

Fie $M_2 = x\,x$.
Obținem $\Gamma_{M_2} = \{x : X\}$ și $\overline{M_2} = M_2$.

$$\frac{\{x : X\} \vdash x : \tau_1 \triangleright C_1 \quad \{x : X\} \vdash x : \tau_2 \triangleright C_2 \quad C_M = C_1 \cup C_2 \cup \{\tau_1 \doteq \tau_2 \to V\}}{\{x : X\} \vdash (x\,x) : V \triangleright C_{M_2}} \ (\to_E^*)$$

$$C_1 = \{\tau_1 \doteq X\}$$
$$C_2 = \{\tau_2 \doteq X\}$$
$$C_{M_2} = \{\tau_1 \doteq X, \tau_2 \doteq X, \tau_1 \doteq \tau_2 \to V\}$$

Constrângerile C_{M_2} nu au "soluție". Ce înseamnă asta?

Constrângerile au "soluție" dacă se pot unifica.

Termeni

Alfabet:

- $\mathcal{F}$ o mulțime de simboluri de funcții de aritate cunoscută
- $\mathcal{V}$ o mulțime numărabilă de variabile
- $\mathcal{F}$ și $\mathcal{V}$ sunt disjuncte

Termeni peste $\mathcal{F}$ si $\mathcal{V}$:

$$t \ ::= \ x \mid f(t_1, \ldots, t_n)$$

- $n \geq 0$
- x este o variabilă
- f este un simbol de funcție de aritate n

Termeni

Notații:

- constante: simboluri de funcții de aritate 0
- $x, y, z, \ldots$ pentru variabile
- $a, b, c, \ldots$ pentru constante
- $f, g, h, \ldots$ pentru simboluri de funcții arbitrare
- $s, t, u, \ldots$ pentru termeni
- $var(t)$ mulțimea variabilelor care apar în t
- ecuații $s \doteq t$ pentru o pereche de termeni
- $Trm_{\mathcal{F},\mathcal{V}}$ mulțimea termenilor peste $\mathcal{F}$ și $\mathcal{V}$

Legătura cu teoria tipurilor

Mulțimea tipurilor simple $\mathbb{T} = \mathbb{V} \mid \mathbb{T} \to \mathbb{T}$

În acest caz, avem alfabetul:

- $\mathcal{F} = \{\to\}$, iar aritatea lui $\to$ este 2
- $\mathcal{V} = \mathbb{V}$

Dacă avem și alte tipuri, extindem $\mathcal{F}$ cu noi simboluri. De exemplu,

- $\mathtt{Unit}, \mathtt{Void}$ cu aritate 0 (deci constante)
- $\mathtt{Bool}, \mathtt{Nat}$ cu aritate 0 (deci constante)
- $\mathtt{Maybe}, \mathtt{List}$ cu aritate 1
- $\times$ cu aritate 2
- $\ldots$

Substituții

O substituție Θ este o funcție de la variabile la termeni,

$$\Theta : \mathcal{V} \to Trm_{\mathcal{F},\mathcal{V}}$$

Notație pentru substituții care schimbă un număr finit de variabile:

$$[u_1/x_1, u_2/x_2, \ldots, u_n/x_n]$$

Aplicarea unei substituții Θ unui termen t:

$$\Theta(t) = \begin{cases} \Theta(x), \text{ dacă } t = x \\ f(\Theta(t_1), \ldots, \Theta(t_n)), \text{ dacă } t = f(t_1, \ldots, t_n) \end{cases}$$

Notație pentru $\Theta(t)$: $t[u_1/x_1, u_2/x_2, \ldots, u_n/x_n]$

Unificare

Doi termeni t_1 și t_2 se unifică dacă există o substituție Θ astfel încât

$$\Theta(t_1) = \Theta(t_2).$$

În acest caz, Θ se numește un unificator al termenilor t_1 și t_2.

Un unificator Θ pentru t_1 și t_2 este cel mai general unificator (cmgu,mgu) dacă pentru orice alt unificator Θ' pentru t_1 și t_2, există o substituție Δ astfel încât

$$\Theta' = \Theta; \Delta.$$

De exemplu, dacă Θ este $[u_1/x_1, u_2/x_2, \ldots, u_n/x_n]$, atunci Δ este de forma $[v_1/y_1, v_2/y_2, \ldots, v_n/y_m]$ cu $x_i \neq y_j$ pentru orice alegere a lui i și j, și

$$\Theta' = [\Delta(u_1)/x_1, \ldots \Delta(u_n)/x_n, v_1/y_1, \ldots, v_m/y_m]$$

Unificatori

Exemplu:

- $t = x + (y * y) = +(x, *(y, y))$
- $t' = x + (y * x) = +(x, *(y, x))$
- $\Theta = \{x \mapsto y\}$
  - $\Theta(t) = y + (y * y)$
  - $\Theta(t') = y + (y * y)$
  - Θ este cmgu
- $\Theta' = \{x \mapsto 0, y \mapsto 0\}$
  - $\Theta'(t) = 0 + (0 * 0)$
  - $\Theta'(t') = 0 + (0 * 0)$
  - $\Theta' = \Theta; \{y \mapsto 0\}$
  - Θ' este unificator, dar nu este cmgu

Algoritmul de unificare

Algoritmul de unificare

- Pentru o mulțime finită de termeni $\{t_1, \ldots, t_n\}$, $n \geq 2$, algoritmul de unificare stabilește dacă există un cmgu.

- Există algoritmi mai eficienți, dar îl alegem pe acesta pentru simplitatea sa.

- Algoritmul lucrează cu două liste:
  - Lista soluție: S
  - Lista de rezolvat: R

- Inițial:
  - Lista soluție: $S = \emptyset$
  - Lista de rezolvat: $R = \{t_1 \doteq t_2, \ldots, t_{n-1} \doteq t_n\}$
    $\doteq$ este un simbol nou care ne ajută să formăm perechi de termeni ("ecuații")

Algoritmul de unificare

Algoritmul constă în aplicarea regulilor de mai jos:

- SCOATE
  - orice ecuație de forma $t \doteq t$ din R este eliminată.
- DESCOMPUNE
  - orice ecuație de forma $f(t_1, \ldots, t_n) \doteq f(t'_1, \ldots, t'_n)$ din R este înlocuită cu ecuațiile $t_1 \doteq t'_1, \ldots, t_n \doteq t'_n$.
- REZOLVĂ
  - orice ecuație de forma $x \doteq t$ sau $t \doteq x$ din R, unde variabila x nu apare în termenul t, este mutată sub forma $x \doteq t$ în S.
    În toate celelalte ecuații (din R și S), x este înlocuit cu t.

Algoritmul de unificare

Algoritmul se termină normal dacă $R = \emptyset$.
În acest caz, S conține cmgu.

Algoritmul este oprit cu concluzia inexistenței unui unificator dacă:

1. În R există o ecuație de forma
$$f(t_1, \ldots, t_n) \doteq g(t'_1, \ldots, t'_k) \text{ cu } f \neq g.$$
2. În R există o ecuație de forma $x \doteq t$ sau $t \doteq x$ și variabila x apare în termenul t.

Algoritmul de unificare - schemă

	Lista soluție S	Lista de rezolvat R
Inițial	$\emptyset$	$t_1 \doteq t'_1, \ldots, t_n \doteq t'_n$
SCOATE	S	R', $t \doteq t$
	S	R'
DESCOMPUNE	S	R', $f(t_1, \ldots, t_n) \doteq f(t'_1, \ldots, t'_n)$
	S	R', $t_1 \doteq t'_1, \ldots t_n \doteq t'_n$
REZOLVĂ	S	R', $x \doteq t$ sau $t \doteq x$, x nu apare în t
	$x \doteq t$, $S[t/x]$	$R'[t/x]$
Final	S	$\emptyset$

$S[t/x]$: în toate ecuațiile din S, x este înlocuit cu t

Exemplul 1

Ecuațiile $\{g(y) \doteq x,\ f(x, h(x), y) \doteq f(g(z), w, z)\}$ au cmgu?

Exemplul 1

Ecuațiile $\{g(y) \doteq x,\ f(x, h(x), y) \doteq f(g(z), w, z)\}$ au cmgu?

S	R	
$\emptyset$	$g(y) \doteq x,\ f(x, h(x), y) \doteq f(g(z), w, z)$	REZOLVĂ
$x \doteq g(y)$	$f(g(y), h(g(y)), y) \doteq f(g(z), w, z)$	DESCOMPUNE
$x \doteq g(y)$	$g(y) \doteq g(z),\ h(g(y)) \doteq w,\ y \doteq z$	REZOLVĂ
$w \doteq h(g(y))$, $x \doteq g(y)$	$g(y) \doteq g(z),\ y \doteq z$	REZOLVĂ
$y \doteq z, x \doteq g(z)$, $w \doteq h(g(z))$	$g(z) \doteq g(z)$	SCOATE
$y \doteq z, x \doteq g(z)$, $w \doteq h(g(z))$	$\emptyset$	

$\Theta = \{y \mapsto z,\ x \mapsto g(z),\ w \mapsto h(g(z))\}$ este cmgu.

Exemplul 2

Ecuațiile $\{g(y) \doteq x,\ f(x, h(y), y) \doteq f(g(z), b, z)\}$ au cmgu?

Exemplul 2

Ecuațiile $\{g(y) \doteq x,\ f(x, h(y), y) \doteq f(g(z), b, z)\}$ au cmgu?

S	R	
$\emptyset$	$g(y) \doteq x,\ f(x, h(y), y) \doteq f(g(z), b, z)$	REZOLVĂ
$x \doteq g(y)$	$f(g(y), h(y), y) \doteq f(g(z), b, z)$	DESCOMPUNE
$x \doteq g(y)$	$g(y) \doteq g(z),\ h(y) \doteq b,\ y \doteq z$	- EȘEC -

- *h* și *b* sunt simboluri de funcții diferite!
- Nu există unificator pentru acești termeni.

Exemplul 3

Ecuațiile $\{g(y) \doteq x,\ f(x, h(x), y) \doteq f(y, w, z)\}$ au cmgu?

Exemplul 3

Ecuațiile $\{g(y) \doteq x,\ f(x, h(x), y) \doteq f(y, w, z)\}$ au cmgu?

S	R	
$\emptyset$	$g(y) \doteq x,\ f(x, h(x), y) \doteq f(y, w, z)$	REZOLVĂ
$x \doteq g(y)$	$f(g(y), h(g(y)), y) \doteq f(y, w, z)$	DESCOMPUNE
$x \doteq g(y)$	$g(y) \doteq y,\ h(g(y)) \doteq w,\ y \doteq z$	- EȘEC -

- În ecuația $g(y) \doteq y$, variabila y apare în termenul $g(y)$.
- Nu există unificator pentru aceste ecuații.

Exemplul 4

Înapoi la constrângerea obținută când am vorbit de *type inference* pentru termenul $M_1 = (\lambda z.\, \lambda u.\, z)\,(y\, x)$.

Am obținut constrângerile

$$C_{M_1} = \{\delta \doteq Z, \tau_1' \doteq (U \to \delta), \tau_1 \doteq (Z \to \tau_1'), \sigma_1 \doteq Y, \sigma_2 \doteq X,$$
$$\sigma_1 \doteq \sigma_2 \to \tau_2, \tau_1 \doteq (\tau_2 \to V)\}$$

- $\to$ simbol de funcție de aritate 2
- $\delta, \tau_1, \tau_1', \tau_2, \sigma_1, \sigma_2, X, Y, Z, U, V$ variabile

Exemplul 4 (cont.)

S	R	
$\emptyset$	$\delta \doteq Z,\ \tau'_1 \doteq (U \to \delta),\ \tau_1 \doteq (Z \to \tau'_1),\ \sigma_1 \doteq Y$ $\sigma_2 \doteq X,\ \sigma_1 \doteq \sigma_2 \to \tau_2,\ \tau_1 \doteq (\tau_2 \to V)$	REZ.
$\delta \doteq Z$	$\tau'_1 \doteq (U \to Z),\ \tau_1 \doteq (Z \to \tau'_1),\ \sigma_1 \doteq Y$ $\sigma_2 \doteq X,\ \sigma_1 \doteq \sigma_2 \to \tau_2,\ \tau_1 \doteq (\tau_2 \to V)$	REZ.
$\delta \doteq Z,\ \tau'_1 \doteq (U \to Z)$	$\tau_1 \doteq (Z \to (U \to Z)),\ \sigma_1 \doteq Y$ $\sigma_2 \doteq X,\ \sigma_1 \doteq \sigma_2 \to \tau_2,\ \tau_1 \doteq (\tau_2 \to V)$	REZ.
$\delta \doteq Z,\ \tau'_1 \doteq (U \to Z),$ $\tau_1 \doteq (Z \to (U \to Z))$	$\sigma_1 \doteq Y,\ \sigma_2 \doteq X,\ \sigma_1 \doteq \sigma_2 \to \tau_2,$ $(Z \to (U \to Z)) \doteq (\tau_2 \to V)$	DESC.
$\delta \doteq Z,\ \tau'_1 \doteq (U \to Z),$ $\tau_1 \doteq (Z \to (U \to Z))$	$\sigma_1 \doteq Y,\ \sigma_2 \doteq X,\ \sigma_1 \doteq \sigma_2 \to \tau_2,$ $Z \doteq \tau_2,\ U \to Z \doteq V$	REZ.
$\delta \doteq Z,\ \tau'_1 \doteq (U \to Z),$ $\tau_1 \doteq (Z \to (U \to Z)),$ $\sigma_1 \doteq Y$	$\sigma_2 \doteq X,\ Y \doteq \sigma_2 \to \tau_2,$ $Z \doteq \tau_2,\ U \to Z \doteq V$	REZ.
$\delta \doteq Z,\ \tau'_1 \doteq (U \to Z),$ $\tau_1 \doteq (Z \to (U \to Z)),$ $\sigma_1 \doteq Y, \sigma_2 \doteq X$	$Y \doteq X \to \tau_2,$ $Z \doteq \tau_2,\ U \to Z \doteq V$	REZ.

Exemplul 4 (cont.)

S	R	
$\delta \doteq Z,\ \tau_1' \doteq (U \to Z),$ $\tau_1 \doteq (Z \to (U \to Z)),$ $\sigma_1 \doteq Y, \sigma_2 \doteq X,\ \tau_2 \doteq Z$	$Y \doteq X \to Z,\ U \to Z \doteq V$	REZ.
$\delta \doteq Z,\ \tau_1' \doteq (U \to Z),$ $\tau_1 \doteq (Z \to (U \to Z)),$ $\sigma_1 \doteq X \to Z, \sigma_2 \doteq X,\ \tau_2 \doteq Z$ $Y \doteq X \to Z$	$U \to Z \doteq V$	REZ.
$\delta \doteq Z,\ \tau_1' \doteq (U \to Z),$ $\tau_1 \doteq (Z \to (U \to Z)),$ $\sigma_1 \doteq X \to Z, \sigma_2 \doteq X,\ \tau_2 \doteq Z$ $Y \doteq X \to Z,\ V \doteq U \to Z$		

Constrângerile se pot unifica!

Exemplul 5

Înapoi la constrângerea obținută când am vorbit de *type inference* pentru termenul $M_2 = x\,x$.

Am obținut constrângerile

$$C_{M_2} = \{\tau_1 \doteq X, \tau_2 \doteq X, \tau_1 \doteq \tau_2 \to V\}$$

- $\to$ simbol de funcție de aritate 2
- τ_1, τ_2, V variabile

Exemplul 5 (cont.)

S	R	
$\emptyset$	$\tau_1 \doteq X,\ \tau_2 \doteq X,\ \tau_1 \doteq \tau_2 \to V$	REZ.
$\tau_1 \doteq X$	$\tau_2 \doteq X,\ X \doteq \tau_2 \to V$	REZ.
$\tau_1 \doteq X,\ \tau_2 \doteq X$	$X \doteq X \to V$	- EȘEC -

- În ecuația $X \doteq X \to V$, variabila X apare în termenul $X \to V$.
- Nu există unificator pentru aceste ecuații.

Exerciții

Considerăm

- x, y, z, u, v variabile,
- a, b, c simboluri de constantă,
- h, g simboluri de funcție de aritate 1,
- f simbol de funcție de aritate 2,
- p simbol de funcție de aritate 3.

Aplicați algoritmul de unificare de mai sus pentru termenii:

1. $p(a, x, h(g(y)))$ și $p(z, h(z), h(u))$
2. $f(h(a), g(x))$ și $f(y, y)$
3. $p(a, x, g(x))$ și $p(a, y, y)$
4. $p(x, y, z)$ și $p(u, f(v, v), u)$

Exerciții - rezolvări

1.

S	R	
$\emptyset$	$p(a, x, h(g(y))) = p(z, h(z), h(u))$	DESCOMPUNE
$\emptyset$	$a \doteq z, x \doteq h(z), h(g(y)) \doteq h(u)$	REZOLVĂ
$z \doteq a$	$x \doteq h(a), h(g(y)) \doteq h(u)$	REZOLVĂ
$z \doteq a, x \doteq h(a)$	$h(g(y)) \doteq h(u)$	DESCOMPUNE
$z \doteq a, x \doteq h(a)$	$g(y) \doteq u$	REZOLVĂ
$z \doteq a, x \doteq h(a), u \doteq g(y)$	$\emptyset$	

$\Theta = \{z/a,\ x/h(a),\ u/g(y)\}$ este cmgu.

Exerciții - rezolvări

2.

S	R	
$\emptyset$	$f(h(a), g(x)) \doteq f(y, y)$	DESCOMPUNE
$\emptyset$	$y \doteq h(a), y \doteq g(x)$	REZOLVĂ
$y \doteq h(a)$	$g(x) \doteq h(a)$	EȘEC

Nu există unificator!

Exerciții - rezolvări

3.

S	R	
$\emptyset$	$p(a, x, g(x)) \doteq p(a, y, y)$	DESCOMPUNE
$\emptyset$	$a \doteq a, x \doteq y, y \doteq g(x)$	SCOATE
$\emptyset$	$x \doteq y, y \doteq g(x)$	REZOLVĂ
$x \doteq y$	$y \doteq g(y)$	EȘEC

Nu există unificator!

Exerciții - rezolvări

4.

S	R	
$\emptyset$	$p(x, y, z) \doteq p(u, f(v, v), u)$	DESCOMPUNE
$\emptyset$	$x \doteq u, y \doteq f(v, v), z \doteq u$	REZOLVĂ
$x \doteq u$	$y \doteq f(v, v), z \doteq u$	REZOLVĂ
$y \doteq f(v, v), x \doteq u$	$z \doteq u$	REZOLVĂ
$z \doteq u, y \doteq f(v, v), x \doteq u$		

$\Theta = \{z/u,\ y/f(v, v),\ x/u\}$ este cmgu.

Pe data viitoare!