P. Urzyczyn: Materiały do wykładu z semantyki

Uproszczony[1] język PCF

Składnia: Poniżej Γ oznacza *otoczenie typowe*, czyli zbiór deklaracji postaci $(x : \tau)$.

$$\Gamma, x:\tau \vdash x : \tau \qquad \Gamma \vdash 0 : \mathsf{int} \qquad \Gamma \vdash \mathit{succ} : \mathsf{int} \to \mathsf{int} \qquad \Gamma \vdash \mathit{pred} : \mathsf{int} \to \mathsf{int}$$

$$\frac{\Gamma, x:\tau \vdash M : \sigma}{\Gamma \vdash \lambda x:\tau.\, M : \tau \to \sigma} \qquad \frac{\Gamma \vdash M : \tau \to \sigma \quad \Gamma \vdash N : \tau}{\Gamma \vdash MN : \sigma}$$

$$\frac{\Gamma \vdash B : \mathsf{int} \quad \Gamma \vdash M : \tau \quad \Gamma \vdash N : \tau}{\Gamma \vdash \mathtt{ifzero}\ B\ \mathtt{then}\ M\ \mathtt{else}\ N : \tau} \qquad \frac{\Gamma, x:\tau \vdash M : \tau}{\Gamma \vdash \mathtt{fix}\ x{:}\tau.\, M : \tau}$$

Skróty: $\Omega_\tau = \mathtt{fix}\ x{:}\tau.\, x$, $\underline{n} = \mathit{succ}^n(0)$. Tam gdzie można pomijamy typy w wyrażeniach.

Semantyka małych kroków:

W języku PCF (w wersji CBN) mamy sześć rodzajów *redeksów*:

$$\begin{array}{rl}
\mathit{pred}(\underline{n+1}) \Rightarrow \underline{n} & \quad \mathit{pred}(0) \Rightarrow 0 \\
(\lambda x.\, M)N \Rightarrow M[N/x] & \quad \mathtt{fix}\ x.\, M \Rightarrow M[\mathtt{fix}\ x.\, M/x] \\
\mathtt{ifzero}\ 0\ \mathtt{then}\ M\ \mathtt{else}\ N \Rightarrow M & \quad \mathtt{ifzero}\ \underline{n+1}\ \mathtt{then}\ M\ \mathtt{else}\ N \Rightarrow N
\end{array}$$

Konteksty ewaluacyjne definiujemy pseudo-gramatyką

$$E[\ ] ::= [\ ] \mid E[\ ]N \mid \mathit{succ}\, E[\ ] \mid \mathit{pred}\, E[\ ] \mid \mathtt{ifzero}\ E[\ ]\ \mathtt{then}\ M\ \mathtt{else}\ N$$

Wartość to liczebnik $\underline{n}$, stała *pred* lub *succ*, lub abstrakcja. Każdy term, który nie jest wartością, można jednoznacznie przedstawić w postaci $E[R]$, gdzie R jest redeksem lub zmienną. (Dowód przez łatwą indukcję.)

Zauważmy, że termy postaci $E[\Omega]$ redukują się same do siebie.

Przyjmujemy $M \to_\ell N$, gdy $M = E[R]$ oraz $N = E[L]$ dla pewnej reguły $R \Rightarrow L$. Jak zwykle, notacja $\twoheadrightarrow_\ell$ oznacza domknięcie przechodnio-zwrotne relacji $\to_\ell$, a notacja $=_\ell$ odpowiednią relację równoważności. Indeks $_\ell$ opuszczamy tam, gdzie można. Napis $\Gamma \vdash M = N : \tau$ oznacza, że $M =_\ell N$ oraz $\Gamma \vdash M : \tau$ i $\Gamma \vdash N : \tau$.

Semantyka dużych kroków (dla termów zamkniętych):

Definiujemy $V \Downarrow V$, gdy V jest wartością, oraz:

$$\frac{M \Downarrow 0}{\mathit{pred}\, M \Downarrow 0} \qquad \frac{M \Downarrow \underline{n+1}}{\mathit{pred}\, M \Downarrow \underline{n}} \qquad \frac{M \Downarrow \underline{n}}{\mathit{succ}\, M \Downarrow \underline{n+1}}$$

$$\frac{M \Downarrow 0 \qquad P \Downarrow V}{\mathtt{ifzero}\ M\ \mathtt{then}\ P\ \mathtt{else}\ Q \Downarrow V} \qquad \frac{M \Downarrow \underline{n+1} \qquad Q \Downarrow V}{\mathtt{ifzero}\ M\ \mathtt{then}\ P\ \mathtt{else}\ Q \Downarrow V}$$

[1]Zwykle dodaje się typ bazowy bool i normalny „`if`”. Można też dodać np. produkt kartezjański itp.

$$\frac{M \Downarrow \lambda x\, A \quad A[N/x] \Downarrow V}{MN \Downarrow V} \qquad\qquad \frac{M[\mathtt{fix}\; x.\, M/x] \Downarrow V}{\mathtt{fix}\; x.\, M \Downarrow V}$$

Semantyka małych kroków jest równoważna semantyce dużych kroków w następującym sensie:

Fakt 1 *Warunki $M \Downarrow V$ i $M \twoheadrightarrow V$ są równoważne.*

Dowód: Z lewej do prawej prosta indukcja ze względu na wyprowadzenie $M \Downarrow V$. Z prawej do lewej indukcja ze względu na liczbę kroków redukcji. Krok indukcyjny wymaga pokazania, że jeśli $M \to M' \Downarrow V$ to $M \Downarrow V$. ∎

Symulacje: Relacja *aplikacyjnej symulacji* $M \sqsubseteq^a N$ zachodzi pomiędzy zamkniętymi termami M i N (tego samego typu), gdy dla dowolnych termów zamkniętych $\vec{N}$:

$$\text{jeśli } M\vec{N} : \mathsf{int}, \text{ oraz } M\vec{N} \Downarrow \underline{n}, \text{ to także } N\vec{N} \Downarrow \underline{n}.$$

Domknięcie quasiporządku $\sqsubseteq^a$ do relacji równoważności oznaczamy przez $\equiv^a$ i nazywamy *równoważnością aplikacyjną*.

Symulacja obserwacyjna $M \sqsubseteq N$ ma miejsce wtedy, gdy dla dowolnego kontekstu $C_{[\ ]}$, z tego, że $C_{[M]} \Downarrow \underline{n}$ wynika $C_{[N]} \Downarrow \underline{n}$. Napis $M \equiv N$ oznacza, że zachodzą obie symulacje $M \sqsubseteq N$ oraz $N \sqsubseteq M$; mówimy wtedy, że M i N są *obserwacyjnie równoważne*.

Fakt 2 (Lemat kontekstowy)
Relacje obserwacyjnej i aplikacyjnej symulacji (równoważności) są takie same.

Dowód: Załóżmy, że $M \sqsubseteq^a N$ i $C_{[M]} \Downarrow \underline{n}$. Dowodzimy, że $C_{[N]} \Downarrow \underline{n}$, przez indukcję ze względu na definicję relacji $\Downarrow$. Są dwa główne przypadki. Pierwszy dotyczy kontekstu z redeksem czołowym, np. $C_{[M]} = (\lambda x\, P_{[M]})Q_{[M]}\vec{R}_{[M]}$. Wtedy możemy zastosować założenie indukcyjne do termu $C'_{[M]} = P_{[M]}[Q_{[M]}/x]\vec{R}_{[M]}$ i otrzymać $P_{[N]}[Q_{[N]}/x]\vec{R}_{[N]} \Downarrow \underline{n}$.

Drugi przypadek jest wtedy, kiedy $C_{[\ ]} = [\ ]P_{[\ ]}\vec{Q}_{[\ ]}$, czyli $C_{[M]} = MP_{[M]}\vec{Q}_M$. Rozpatrzmy kontekst $D_{[\ ]} = MP_{[\ ]}\vec{Q}_{[\ ]}$. Wtedy $D_{[M]} = C_{[M]}$ oraz $D_{[N]} = MP_{[N]}\vec{Q}_{[N]}$. Ponieważ to nie jest liczebnik, więc kontekst $D_{[\ ]}$ musi mieć redeks czołowy i stosuje się do niego poprzedni przypadek. Zatem $D_{[N]} \Downarrow \underline{n}$. Teraz korzystamy z tego, że $M \sqsubseteq^a N$ i dostajemy $C_{[N]} \Downarrow \underline{n}$. Pozostałe przypadki są łatwe. ∎

Wniosek 3 *Dla każdego M zachodzi $\Omega \sqsubseteq M$, bo zawsze $\Omega\vec{N} \not\Downarrow$.*

Semantyka denotacyjna: Dziedziny semantyczne to $D_{\mathsf{int}} = \mathbb{N}_\perp$, oraz $D_{\tau\to\sigma} = [D_\tau \to D_\sigma]$. Poniżej zakładamy, że $\rho(x) \in D_\tau$, gdy $(x : \tau) \in \Gamma$.

$$[\![x]\!]_\rho = \rho(x), \qquad [\![0]\!]_\rho = 0, \qquad [\![\mathit{succ}]\!]_\rho(n) = n+1, \qquad [\![\mathit{succ}]\!]_\rho(\perp) = \perp$$

$$[\![\mathit{pred}]\!]_\rho(n+1) = n, \qquad [\![\mathit{pred}]\!]_\rho(0) = 0, \qquad [\![\mathit{pred}]\!]_\rho(\perp) = \perp$$

$$[\![MN]\!]_\rho = [\![M]\!]_\rho([\![N]\!]_\rho), \qquad [\![\lambda x.\, M]\!]_\rho(d) = [\![M]\!]_{\rho[x\mapsto d]}$$

$$[\![\mathtt{fix}\ x{:}\tau.\, M]\!]_\rho = \textsc{lfp}(\lambda\!\!\lambda d.\, [\![M]\!]_{\rho[x\mapsto d]})$$

$$[\![\mathtt{ifzero}\ B\ \mathtt{then}\ M\ \mathtt{else}\ N]\!]_\rho = \begin{cases} \bot, & \text{jeśli } [\![B]\!]_\rho = \bot; \\ [\![M]\!]_\rho, & \text{jeśli } [\![B]\!]_\rho = 0; \\ [\![N]\!]_\rho, & \text{w przeciwnym przypadku.} \end{cases}$$

Fakt 4 (Poprawność) *Jeśli* $M =_\ell N$, *to* $[\![M]\!] = [\![N]\!]$.

Z powyższego wynika w szczególności, że jeśli $M \Downarrow \underline{n}$, to $[\![M]\!] = n$.

Aproksymowanie punktu stałego: Dla dowolnego M definiujemy

$$\mathtt{fix}^0 x.\, M = \Omega, \qquad \mathtt{fix}^{k+1} x.\, M = M[\mathtt{fix}^k x.\, M/x].$$

Lemat 5 *Dla dowolnego* M *i dowolnego* k*:*

1. $\mathtt{fix}^k x.\, M \sqsubseteq \mathtt{fix}\ x.\, M$;
2. $[\![\mathtt{fix}\ x.\, M]\!] = \sup_k [\![\mathtt{fix}^k x.\, M]\!]$.

Dowód: (1) Indukcja ze względu na k. Korzystamy z tego, że zawsze $\Omega \sqsubseteq N$, i z tego, że $N_1 \sqsubseteq N_2$ implikuje $C[N_1] \sqsubseteq C[N_2]$. (2) Natychmiast z tw. o punkcie stałym. ∎

Termy obliczalne: Dla dowolnego τ definiujemy zbiór termów zamkniętych $[\tau]$ typu τ:

- $[\mathsf{int}] = \{M : \mathsf{int} \mid M \Downarrow \text{ lub } [\![M]\!] = \bot\}$;
- $[\tau \to \sigma] = \{M : \tau \to \sigma \mid \forall N (N \in [\tau] \to MN \in [\sigma])\}$.

Elementy zbioru $[\tau]$ nazywamy termami *obliczalnymi* w τ. Jeśli $x : \tau_1, \ldots, x : \tau_n \vdash M : \tau$, to M jest *obliczalny*, gdy dla dowolnych $N_1 \in [\tau_1], \ldots, N_n \in [\tau_n]$ zachodzi $M[N_1/x_1 \ldots N_n/x_n] \in [\tau]$. Term $M[N_1/x_1 \ldots N_n/x_n] \in [\tau]$ nazwiemy *obliczalną instancją* termu M.

Lemat 6 *Term* M *jest obliczalny wtedy i tylko wtedy, gdy dla dowolnej obliczalnej instancji* M' *termu* M *i dowolnych obliczalnych i zamkniętych termów* $\vec{N}$*:*

$$\textit{jeśli } M'\vec{N} : \mathbf{int}, \textit{ to albo } [\![M'\vec{N}]\!] = \bot \textit{ albo } M'\vec{N} \Downarrow.$$

Fakt 7 (Słaba adekwatność) *Jeśli* $[\![M]\!] = n$ *to* $M \Downarrow \underline{n}$.

Dowód: Wystarczy udowodnić, że każdy term jest obliczalny. Dowód (opuszczony) jest przez indukcję ze względu na budowę termu. W przypadku termu postaci $\mathtt{fix}\ x.\, M$, korzystamy z równości $[\![\mathtt{fix}\ x.\, M]\!] = \sup_k [\![\mathtt{fix}^k x.\, M]\!]$. ∎

Wniosek 8 (Adekwatność) *Jeśli* $[\![M]\!] = [\![N]\!]$ *to* M *i* N *są obserwacyjnie równoważne.*

Dowód: Na mocy „lematu kontekstowego” wystarczy wykazać równoważność aplikacyjną. Ta wynika za słabej adekwatności, bo $[\![M\vec{P}]\!] = [\![N\vec{P}]\!]$ dla dowolnych $\vec{P}$. ∎

Równoległa alternatywa: Funkcja *por* : $\mathbb{N}_\perp \to \mathbb{N}_\perp \to \mathbb{N}_\perp$ jest określona tak:

$$\mathit{por}\, m\, n = \begin{cases} 0, & \text{jeśli } m = 0 \text{ lub } n = 0; \\ 1, & \text{jeśli } m, n \in \mathbb{N} \text{ i } m, n > 0; \\ \perp, & \text{w przeciwnym przypadku.} \end{cases}$$

Przyjmijmy, że zero interpretujemy jako *true* a liczby większe od zera jako *false*. Wtedy funkcja *por* to *równoległa alternatywa* („parallel or”), spełniona zawsze gdy któryś ze składników jest spełniony, nawet wtedy, gdy nieokreślony jest drugi składnik.

Lemat 9 *Załóżmy, że* $M' = M[\vec{N}/\vec{x}] \twoheadrightarrow \underline{n}$*, gdzie* $\vec{x}$ *to wszystkie zmienne wolne w* M*, a termy* $\vec{N}$ *są zamknięte. Wtedy albo* $M \twoheadrightarrow \underline{n}$*, albo* $M \twoheadrightarrow E[x]$*, gdzie* $x \in \vec{x}$ *i* $E[\ ]$ *jest kontekstem ewaluacyjnym.*

Dowód: Indukcja ze względu na długość redukcji $M' \twoheadrightarrow \underline{n}$. Zakładając, że $M = N\vec{R}$, gdzie term N nie jest aplikacją, mamy tu kilka możliwości w zależności od N. Zbadanie ich pozostawiamy czytelnikowi. ∎

Funkcja *por* nie jest definiowalna w języku PCF, co wynika z poniższego lematu (dla $m = 1$).

Lemat 10 *Nie istnieje term* M *o zmiennych wolnych* x *i* y*, spełniający jednocześnie warunki:*

$$M[\underline{0}/x, \Omega/y] =_\ell \underline{0} = M[\Omega/x, \underline{0}/y] \quad \textit{oraz} \quad M[\underline{1}/x, \underline{1}/y] =_\ell \underline{m}, \textit{ gdzie } m > 0.$$

Dowód: Z lematu 9 wynika, że wtedy $M \twoheadrightarrow E[x]$ lub $M \twoheadrightarrow E[y]$, niemożliwe jest bowiem, aby term M redukował się i do $\underline{0}$ i do $\underline{m}$. Jeśli jednak $M \twoheadrightarrow E[x]$, to $M[\Omega/x, \underline{0}/y] \twoheadrightarrow E[\Omega]\not\Downarrow$, i podobnie w drugim przypadku. ∎

Stąd wynika, że nie zachodzi twierdzenie odwrotne do adekwatności.

Fakt 11 *Semantyka denotacyjna dla PCF nie jest w pełni abstrakcyjna: istnieją takie termy zamknięte* T_0 *i* T_1*, że* $T_0 \equiv T_1$ *ale* $[\![T]\!]_0 \neq [\![T]\!]_1$.

Dowód: Niech T_a, gdzie $a \in \{\underline{0}, \underline{1}\}$, oznacza term:

$$\lambda p.\, \texttt{ifzero}\; p\, 0\, \Omega\; \texttt{then}\; \texttt{ifzero}\; p\, \Omega\, 0\; \texttt{then}\; \texttt{ifzero}\; p\, 1\, 1\; \texttt{then}\; \Omega\; \texttt{else}\; a\; \texttt{else}\; \Omega\; \texttt{else}\; \Omega.$$

Łatwo widzieć, że $[\![T_0]\!]\mathit{por} = 0$ i $[\![T_1]\!]\mathit{por} = 1$, skąd $[\![T_0]\!] \neq [\![T_1]\!]$. Ponadto, gdyby $T_0 M\Downarrow$, dla jakiegoś M, to term Mxy musiałby spełniać warunki, o których mowa w lemacie 10. Skoro takiego nie ma, to zawsze $T_0 M\not\Downarrow$, czyli T_0 jest aplikacyjnie równoważne z Ω. Podobnie jest dla T_0, więc mamy aplikacyjną równoważność termów T_0 i T_1. A wiadomo, że z równoważności aplikacyjnej wynika równoważność obserwacyjna. ∎

Ćwiczenia:

1. Udowodnić, że dla każdej funkcji częściowo rekurencyjnej φ istnieje taki term Φ w języku PCF, że $\varphi(n) = m$ wtedy i tylko wtedy, gdy $\Phi(\underline{n}) \Downarrow \underline{m}$. Inaczej: w PCF można zdefiniować każdą funkcję częściowo rekurencyjną.
2. Uzupełnić brakujące dowody.