P. Urzyczyn: Materiały do wykładu z semantyki

Model $\mathcal{D}_\infty$

Ponieważ model $\mathcal{D}_\infty$ jest izomorficzny z przestrzenią funkcji ciągłych $[\mathcal{D}_\infty \to \mathcal{D}_\infty]$ (a zatem także z przestrzenią $[[\mathcal{D}_\infty \to \mathcal{D}_\infty] \to [\mathcal{D}_\infty \to \mathcal{D}_\infty]]$), więc elementy $\mathcal{D}_\infty$ można utożsamiać z funkcjami ciągłymi z $\mathcal{D}_\infty$ do $\mathcal{D}_\infty$, a nawet z $[\mathcal{D}_\infty \to \mathcal{D}_\infty]$ do $[\mathcal{D}_\infty \to \mathcal{D}_\infty]$. Powoduje to schizofreniczną dwuznaczność notacyjną, na szczęście nie prowadzącą do żadnej konfuzji. Jeśli bowiem funkcję $f : \mathcal{D}_\infty \to \mathcal{D}_\infty$ utożsamimy z elementem $f' = G(f)$, to zachodzą równości $f(a) = F(G(f))(a) = F(f')(a) = f' \cdot a$. A więc wartość funkcji f na argumencie a to to samo co wynik działania $f' \cdot a$ w modelu (który zwykle zapiszemy po prostu jako $f \cdot a$). Podobnie, jeśli φ jest funkcją ciągłą z $[\mathcal{D}_\infty \to \mathcal{D}_\infty]$ do $[\mathcal{D}_\infty \to \mathcal{D}_\infty]$, a f jest jak poprzednio, to możemy utworzyć aż trzy różne aplikacje φ do f. Jedna to $\varphi(f) \in [\mathcal{D}_\infty \to \mathcal{D}_\infty]$, druga to $\varphi'(f') \in \mathcal{D}_\infty$, gdzie $\varphi' = G \circ \varphi \circ F$, i wreszcie trzecia to $G(\varphi') \cdot f'$. I znowu wszystkie trzy oznaczają (z dokładnością do odpowiednich utożsamień) ten sam element modelu, bo np. $\varphi'(f') = G(\varphi(F(G(f)))) = G(\varphi(f))$.

0.1 Kombinator Y w modelu $\mathcal{D}_\infty$

Wartością termu $\mathbf{B}' = \lambda xyz.y(xz)$ w modelu $\mathcal{D}_\infty$ jest element $\mathbf{b}' = [\![\mathbf{B}']\!]$, spełniający dla dowolnych $a, b, c \in \mathcal{D}_\infty$ warunek $\mathbf{b}' \cdot a \cdot b \cdot c = b \cdot (a \cdot c)$. Element ten, zgodnie z ogólną schizofrenią, utożsamiamy z funkcją $\mathbf{b}' : \mathcal{D}_\infty \to \mathcal{D}_\infty$ o własności $\mathbf{b}'(a) \cdot b \cdot c = b \cdot (a \cdot c)$. Niech teraz $\mathrm{id} = [\![\mathbf{I}]\!]$. Wtedy $\mathrm{id}_0 = \bot$, oraz $\mathrm{id}_{n+1} = \mathrm{id}_{D_n}$, dla $n \in \mathbb{N}$.

Lemat 0.1 *Element* $\mathrm{id} \in \mathcal{D}_\infty$ *jest jedynym punktem stałym funkcji* $\mathbf{b}'$.

Dowód: Nietrudno zauważyć, że id jest punktem stałym $\mathbf{b}'$, pozostaje pokazać, że $\mathbf{b}'$ ma tylko jeden punkt stały. Załóżmy więc, że ε jest punktem stałym, tj. że $\varepsilon \cdot b \cdot c = b \cdot (\varepsilon \cdot c)$, dla dowolnych b i c. Przypuśćmy, że $x \in D_{n+1}$ i $y \in D_n$ i policzmy co to jest $(\varepsilon \cdot x \cdot y)_n$. Otóż z lematu 11.6(2) wynika, że $(\varepsilon \cdot x \cdot y)_n = (\varepsilon \cdot x)_{n+1}(y) = \varepsilon_{n+2}(x)(y)$. Z drugiej strony mamy $(\varepsilon \cdot x \cdot y)_n = (x \cdot (\varepsilon \cdot y))_n$, bo to punkt stały, a ponieważ z lematu 11.6(1) wynika $x \cdot (\varepsilon \cdot y) = x((\varepsilon \cdot y)_n)$, więc wyrażenie $(x \cdot (\varepsilon \cdot y))_n$ można zapisać jako $(x((\varepsilon \cdot y)_n))_n$. To się dalej upraszcza do $x((\varepsilon \cdot y)_n)$, bo $x((\varepsilon \cdot y)_n) \in D_n$. Ale $(\varepsilon \cdot y)_n = \varepsilon_{n+1}(y)$, bo $y \in D_n$, więc w końcu wychodzi równość $(\varepsilon \cdot x \cdot y)_n = x(\varepsilon_{n+1}(y))$. Mamy więc dla wszystkich $n \in \mathbb{N}$:

$$\varepsilon_{n+2}(x)(y) = x(\varepsilon_{n+1}(y)). \tag{*}$$

Teraz przez indukcję ze względu na n pokażemy, że $\varepsilon_n = \mathrm{id}_n$. Dla $n = 0$ i dowolnego $d \in D_0$, podstawiając $x = \lambda\!\!\lambda z.d$ i $y = \bot$ do (*), dostajemy równość $\varepsilon_2(\lambda\!\!\lambda z.d)(\bot) = d$, której lewa strona to nie co innego niż $\psi_1(\varepsilon_2)(d) = \varepsilon_1(d)$. A więc po prostu $\varepsilon_1(d) = d$ czyli $\varepsilon_1 = \mathrm{id}_1$. Dalej mamy już łatwą indukcję. Jeśli $\varepsilon_{n+1} = \mathrm{id}_{n+1}$ to $\varepsilon_{n+2}(x)(y) = x(y)$, czyli $\varepsilon_{n+2}(x) = x$. Pozostaje jeszcze zauważyć, że $\varepsilon_0 = \varepsilon_1(\bot) = \mathrm{id}_1(\bot) = \bot$. ∎

Niech teraz $f \in [\mathcal{D}_\infty \to \mathcal{D}_\infty]$ i niech $\rho(y) = f$ (pamiętamy o utożsamieniu $f = G(f) \in \mathcal{D}_\infty$). Przyjmijmy oznaczenie $\Delta_f = [\![\lambda x.y(xx)]\!]_\rho$. Wtedy $\Delta_f \cdot a = f(a \cdot a)$, dla każdego a. Przez $\mathrm{lfp}(f)$ oznaczymy najmniejszy punkt stały funkcji f.

Lemat 0.2 *Jeśli* $e \le \mathrm{id}$ *oraz* $e \cdot \Delta_f \cdot \Delta_f \le \mathrm{lfp}(f)$, *to też* $\mathbf{b}'(e) \le \mathrm{id}$ *oraz* $\mathbf{b}'(e) \cdot \Delta_f \cdot \Delta_f \le \mathrm{lfp}(f)$.

Dowód: Liczymy najpierw tak: $\mathbf{b}'(e)(x)(y) = x \cdot (e \cdot y) \le x \cdot y = x(y)$, a potem tak: $\mathbf{b}'(e) \cdot \Delta_f \cdot \Delta_f = \Delta_f \cdot (e \cdot \Delta_f) = f(e \cdot \Delta_f \cdot (e \cdot \Delta_f)) \le f(e \cdot \Delta_f \cdot \Delta_f) \le f(\mathrm{lfp}(f)) = \mathrm{lfp}(f)$. ∎

Wniosek 0.3 *W modelu* $\mathcal{D}_\infty$ *znaczeniem kombinatora punktu stałego* $\mathbf{Y}$ *jest operator najmniejszego punktu stałego* $\lambda\!\!\lambda f.\mathrm{lfp}(f)$. *Ściślej,* $[\![\mathbf{Y}]\!] \cdot a = G(\mathrm{lfp}(F(a)))$, *dla* $a \in \mathcal{D}_\infty$.

Dowód: Niech $e^0 = \bot$ i $e^{n+1} = \mathbf{b}'(e^n)$, dla $n \in \mathbb{N}$. Na mocy lematu 0.2 mamy dla wszystkich n nierówność $e^n \cdot \Delta_f \cdot \Delta_f \le \mathrm{lfp}(f)$, skąd na mocy ciągłości aplikacji wynika, że także $\mathrm{id} = \mathrm{lfp}(\mathbf{b}')$ spełnia warunek $\mathrm{id} \cdot \Delta_f \cdot \Delta_f \le \mathrm{lfp}(f)$. Inaczej mówiąc, $[\![\mathbf{Y}]\!] \cdot f = \Delta_f \cdot \Delta_f \le \mathrm{lfp}(f)$. Oczywiście $[\![\mathbf{Y}]\!] \cdot f$ jest punktem stałym, bo jesteśmy w lambda-modelu. ∎

0.2 Pełna abstrakcyjność modelu $\mathcal{D}_\infty$

Naszkicujemy teraz schemat dowodu twierdzenia 11.11, pomijając co trudniejsze jego fragmenty. Posłużymy się drzewami Böhma, dowodząc w istocie „przy okazji” twierdzenia 8.4. Razem dostaniemy to:

Twierdzenie 0.4 *Następujące warunki są równoważne:*

1. *Termy* M *i* N *są obserwacyjnie równoważne;*
2. $BT(M) \approx_\eta BT(N)$;
3. $\mathcal{D}_\infty \models M = N$.

Implikacja (1) ⇒ (2) to w istocie nieznaczne uogólnienie znanego nam twierdzenia Böhma 8.1. Implikację (3) ⇒ (1) nazywamy *adekwatnością* modelu, a implikacja (1) ⇒ (3) nosi nazwę *pełnej abstrakcyjności*. Adekwatność można uważać za „słabą pełność” (bo nie twierdzimy, że $M =_\beta N$ a tylko, że $M \equiv N$), natomiast pełna abstrakcyjność to „mocna poprawność”, stwierdzająca, że termy nieodróżnialne obliczeniowo (nawet niekoniecznie równe) są interpretowane w ten sam sposób.

Dowód twierdzenia 0.4 wymaga pewnych pojęć pomocniczych. Wprowadzimy dwie relacje dla drzew Böhma. Napis $B \sqsubseteq B'$ oznacza, że B' powstaje z B przez wstawienie jakichś poddrzew w miejsca, w których w B występuje Ω. Relacja $B \preceq_\eta B'$ zachodzi zaś, gdy istnieje (skończony lub nieskończony) ciąg eta-ekspansji $B = B_0 \;{}_\eta\!\!\leftarrow B_1 \;{}_\eta\!\!\leftarrow B_2 \;{}_\eta\!\!\leftarrow B_3 \;{}_\eta\!\!\leftarrow \cdots$ zbieżny do B'. Oczywiście $B \approx_\eta B'$ oznacza, że $B \preceq_\eta B'' \;{}_\eta\!\!\succeq B'$ dla pewnego B''. Notację $\preceq_\eta$ i $\approx_\eta$ stosujemy też dla termów; mamy wtedy na myśli ich drzewa Böhma.

Aproksymant to skończone drzewo Böhma (inaczej: term w postaci normalnej, w którym może występować stała Ω). Dla dowolnego termu M, mamy zbiór *aproksymantów termu* M:

$$A(M) = \{A \mid A \text{ jest aproksymantem oraz } A \sqsubseteq M\}$$

W analogiczny sposób definiujemy $A(T)$ dla dowolnego drzewa T. Przyjmując, że znaczeniem stałej Ω jest $\bot$, możemy przypisać każdemu aproksymantowi wartość w modelu $\mathcal{D}_\infty$. Dalej możemy przyjąć definicję $[\![T]\!]_\rho = \sup\{[\![A]\!]_\rho \mid A \in A(T)\}$.

Najważniejszy fakt, z którego wszystko wynika (w szczególności równość $[\![M]\!]_\rho = [\![BT(M)]\!]_\rho$), i którego dowód teraz opuścimy, to następujące twierdzenie o aproksymacji:

Twierdzenie 0.5 (o aproksymacji) *Wartość dowolnego termu jest kresem górnym wartości jego aproksymantów, tj.* $[\![M]\!]_\rho = \sup\{[\![A]\!]_\rho \mid A \in A(M)\}$.

Lemat 0.6 *Term M jest rozwiązalny wtedy i tylko wtedy, gdy $[\![M]\!]_\rho \neq \bot$ przy pewnym ρ.*

Dowód: Najpierw zauważmy, że jeśli A jest aproksymantem, to $[\![A]\!]_\rho \neq \bot$, dla pewnego ρ, gdy tylko $A \neq \Omega$. Jeśli bowiem $A = \lambda\vec{x}.y\vec{B}$, gdzie y jest zmienną wolną, to wystarczy wziąć $\rho(y) = \lambda\!\!\lambda\vec{a}.d$, gdzie $d \neq \bot$. A jeśli $A = \lambda\vec{x}.x_i\vec{B}$, to należy użyć $\lambda\!\!\lambda\vec{a}.d$ jako i-tego argumentu.

Ponieważ $[\![M]\!]_\rho = \sup\{[\![A]\!]_\rho \mid A \in A(M)\}$, więc z powyższego wynika $[\![M]\!]_\rho \neq \bot$, jeśli tylko M ma choć jeden nietrywialny aproksymant, tj. gdy $BT(M) \neq \Omega$. Przypomnijmy zaś, że term jest rozwiązalny wtedy i tylko wtedy, gdy ma nietrywialne drzewo Böhma (istnieje czołowa postać normalna). ∎

Wniosek 0.7 (adekwatność) *Jeśli $[\![M]\!]_\rho = [\![N]\!]_\rho$ dla dowolnego ρ, to $M \equiv N$.*

Dowód: Jeśli $[\![M]\!]_\rho = [\![N]\!]_\rho$ to także $[\![C[M]]\!]_\rho = [\![C[N]]\!]_\rho$ dla dowolnego kontekstu $C[\ ]$. Jeśli więc $[\![C[M]]\!]_\rho \neq \bot$ to i $[\![C[N]]\!]_\rho \neq \bot$, czyli z rozwiązalności $C[M]$ wynika rozwiązalność $C[N]$. ∎

Teraz udowodnimy pełną abstrakcyjność modelu $\mathcal{D}_\infty$. Na początek taki prosty lemat:

Lemat 0.8 *Niech $T_1 \preceq_\eta T_2$. Wówczas:*

1. *Dla dowolnego $A \in A(T_1)$ istnieje takie $B \in A(T_2)$, że $B \twoheadrightarrow_\eta A$.*
2. *Dla dowolnego $B \in A(T_2)$ istnieje takie $A \in A(T_1)$, że $B \twoheadrightarrow_\eta A$.*

Dowód: (1) W nieskończonym ciągu eta-ekspansji od T_1 do T_2, tylko skończenie wiele kroków dotyczy wierzchołków należących do aproksymanta A. Te eta-ekspansje przekształcają A w pewnego aproksymanta drzewa T_2.

(2) Analogicznie, tylko skończenie wiele eta-ekspansji pozostawia swój ślad w drzewie B. Odpowiadają one pewnej eta-redukcji z B do jakiegoś aproksymanta drzewa T_1. ∎

Lemat 0.9 *Jeśli $T_1 \preceq_\eta T_2$ to $[\![T_1]\!]_\rho = [\![T_2]\!]_\rho$, dla każdego ρ.*

Dowód: Przypomnijmy, że $[\![T_1]\!]_\rho = \sup\{[\![A]\!]_\rho \mid A \in A(M)\}$ i podobnie dla T_2. Ponieważ eta-równe termy są równe w modelu, więc z lematu 0.8 wynika od razu, że odpowiednie kresy są równe. ∎

Stąd natychmiast otrzymujemy:

Wniosek 0.10 (pełna abstrakcyjność) *Jeśli $M \equiv N$, to $[\![M]\!]_\rho = [\![N]\!]_\rho$ dla dowolnego ρ.*

Dowód: Skoro $M \equiv N$, to $BT(M) \approx_\eta BT(N)$, czyli $BT(M) \preceq_\eta T \ {}_\eta{\succeq} BT(N)$. Z lematu 0.9 wynika od razu, że $[\![BT(M)]\!]_\rho = [\![BT(N)]\!]_\rho$. A z twierdzenia o aproksymacji otrzymujemy, że także wartości $[\![M]\!]_\rho$ i $[\![N]\!]_\rho$ muszą być równe. ∎