Notatki do wykładu
Automaty a rekurencja:
Schematy rekurencyjne

Robert Błaszkiewicz

11.10.2016

Niech o będzie typem bazowym. Powiemy, że α jest typem, jeżeli: α jest typem bazowym ($\alpha = o$) lub $\alpha = \beta \to \gamma$ dla β, γ będących typami. Przykłady: $o \to o, (o \to o) \to o, o \to (o \to o), (o \to o) \to (o \to o)$. Uwaga: domyślnie stosujemy nawiasowanie prawostronne tzn. $\alpha \to \beta \to \gamma = (\alpha \to (\beta \to \gamma))$, gdzie α, β, γ są typami. Zdefiniujmy rząd typu ord w następujący sposób: $ord(o) = 0$, $ord(\alpha \to \beta) = max(ord(\alpha) + 1, ord(\beta))$. W szczególności: $ord(\alpha_1 \to \cdots \to \alpha_k \to o) = \max\limits_{i=1\ldots k}(\alpha_i) + 1$ gdy $k > 0$.

Będziemy rozważać gramatyki generujące drzewa etykietowane literami ze skończonego alfabetu. Formalnie, schematem rekurencyjnym (gramatyką) nazwiemy krotkę $\mathcal{S} = ((\Sigma, rank : \Sigma \to \mathbb{N}), \mathcal{N}, S^o \in \mathcal{N}, \mathcal{R})$, gdzie: $(\Sigma, rank)$ to urangowany alfabet (np. $\Sigma = \{a, b, c\}$, $rank = \{(a, 0), (b, 2), (c, 0)\}$), gdzie ranga litery będzie odpowiadać liczbie dzieci wierzchołka etykietowanego daną literą, $\mathcal{N}$ to zbiór otypowanych nieterminali (np. $\mathcal{N} = \{A^{o \to o}, B^{o \to (o \to o) \to o}, S^o\}$), S^o to nieterminal startowy, $\mathcal{R}$ to zbiór reguł.

Pozostaje zdefiniować dozwoloną postać reguł w zbiorze $\mathcal{R}$. Ustalmy urangowany alfabet $(\Sigma, rank : \Sigma \to \mathbb{N})$, zbiór nieterminali $\mathcal{N}$ oraz dodatkowo zbiór otypowanych zmiennych $\mathcal{V}$ (np. $x^{o \to o \to o}, y^{o \to o}, z^o$). Uwaga: zarówno w przypadku zmiennych i nieterminali, jak i termów (zdefiniowanych poniżej), będziemy zwyczajowo umieszczać adnotacje o ich typie w przypisie górnym, jednak, dla zwiększenia czytelności, często przypis ten będziemy pomijać.

Termem nad $((\Sigma, rank), \mathcal{N}, \mathcal{V})$ nazwiemy:

- x^α, gdzie $x^\alpha \in \mathcal{V}$,
- A^α, gdzie $A^\alpha \in \mathcal{N}$,
- $(a\, N_1^o \ldots N_k^o)$, gdzie $N_1^o, \ldots, N_k^o$ – termy nad $((\Sigma, rank), \mathcal{N}, \mathcal{V})$, $a \in \Sigma$, $k = rank(a)$,
- $(N^{\alpha \to \beta}\, M^\alpha)^\beta$, gdzie $N^{\alpha \to \beta}, M^\alpha$ – termy nad $((\Sigma, rank), \mathcal{N}, \mathcal{V})$.

Zbiór $\mathcal{R}$ może zawierać reguły postaci: $A^{\alpha_1 \to \cdots \to \alpha_k \to o}\, x_1^{\alpha_1} \ldots x_k^{\alpha_k} \longrightarrow M^o$, gdzie $A^{\alpha_1 \to \cdots \to \alpha_k \to o} \in \mathcal{N}$, $x_1^{\alpha_1}, \ldots, x^{\alpha_k}$ – otypowane zmienne, M^o – term nad $((\Sigma, rank), \mathcal{N}, \{x_1^{\alpha_1}, \ldots, x_k^{\alpha_k}\})$.

W dalszej części przyjmiemy, że mamy zadany schemat rekurencyjny $\mathcal{S} = ((\Sigma, rank), \mathcal{N}, S^o, \mathcal{R})$. Zdefiniujmy relację $\longrightarrow_{\mathcal{S}}$ jako najmniejszą relację spełniającą następujące warunki:

- $A\, N_1 \ldots N_k \longrightarrow_{\mathcal{S}} M[N_1/x_1, \ldots, N_k/x_k]$ jeśli $(A\, x_1 \ldots x_k \longrightarrow M) \in \mathcal{R}$,
- $a\, N_1 \ldots N_{i-1}\, N_i\, N_{i+1} \ldots N_k \longrightarrow_{\mathcal{S}} a\, N_1 \ldots N_{i-1}\, N_i'\, N_{i+1} \ldots N_k$ jeśli $N_i \longrightarrow_{\mathcal{S}} N_i'$ dla pewnego $i \in \{1 \ldots k\}$.

Przez $\longrightarrow^*_{\mathcal{S}}$ będziemy oznaczać domknięcie przechodnio-zwrotne tej relacji.

Drzewo generowane z termu N otrzymujemy w następujący sposób:

1. jeśli $N \longrightarrow^*_{\mathcal{S}} a\, N_1 \ldots N_k$ dla jakichś termów $N_1, \ldots, N_k \in \mathcal{N}$ i $a \in \Sigma$, to drzewo ma w korzeniu a, a i-te poddrzewo to drzewo generowane z N_i,
2. w.p.p. w korzeniu umieszczamy specjalną etykietę ω ($\omega \notin \Sigma$).

Przykład 1.

$(\Sigma, rank) = (\{a, b, c\}, \{(a, 0), (b, 2), (c, 0)\})$

$\mathcal{N} = \{S^o, H^{o\to o}, F^{(o\to o)\to o}, B^{o\to o\to o}, C^{o\to o}\}$

Zbiór reguł ($\mathcal{R}$):

$S^o \longrightarrow (H^{o\to o}\, a^o)^o$

$H^{o\to o}\, x^o \longrightarrow (F^{(o\to o)\to o}\, (B^{o\to o\to o}\, x^o)^{o\to o})^o$

$F^{(o\to o)\to o}\, \phi^{o\to o} \longrightarrow (\phi^{o\to o}\, (\phi^{o\to o}\, (F^{(o\to o)\to o}\, C^{o\to o})^o)^o)^o$

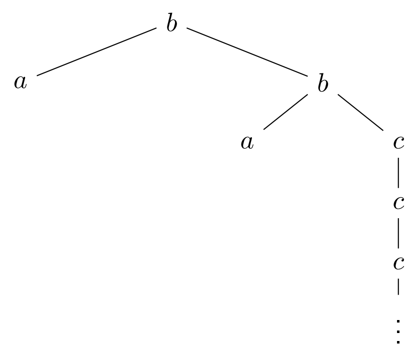
$B^{o\to o\to o}\, x^o\, y^o \longrightarrow (b\, x^o\, y^o)^o$

$C^{o\to o}\, x^o \longrightarrow (c\, x^o)^o$

Zaczynając od symbolu startowego S i stosując nasze reguły otrzymujemy:

$S \longrightarrow_{\mathcal{S}} H\,a \longrightarrow_{\mathcal{S}} F\,(B\,a) \longrightarrow_{\mathcal{S}} B\,a\,(B\,a\,(F\,C)) \longrightarrow_{\mathcal{S}} b\,a\,(B\,a\,(F\,C)) \longrightarrow_{\mathcal{S}} b\,a\,(b\,a\,(F\,C)) \longrightarrow_{\mathcal{S}} b\,a\,(b\,a\,(C\,(C\,(F\,C)))) \longrightarrow_{\mathcal{S}} b\,a\,(b\,a\,(c\,(C\,(F\,C)))) \longrightarrow_{\mathcal{S}} b\,a\,(b\,a\,(c\,(c\,(F\,C)))) \longrightarrow_{\mathcal{S}} b\,a\,(b\,a\,(c\,(c\,(C\,(C\,(F\,C)))))) \longrightarrow_{\mathcal{S}} \ldots$

Łatwo zauważyć, że generowane jest następujące drzewo:



Przykład 2.

$(\Sigma, rank) = (\{a\}, \{(a, 1)\})$

$\mathcal{N} = \{S^o, T^o\}$

Zbiór reguły ($\mathcal{R}$):

$S \longrightarrow a\,T$

$T \longrightarrow T$

Otrzymujemy:

$S \longrightarrow_{\mathcal{S}} a\,T \longrightarrow_{\mathcal{S}} a\,T \longrightarrow_{\mathcal{S}} \ldots$

Wyżej wymienione przykłady pokazują schematy deterministyczne, tzn. takie, w których dla każdego nieterminala jest dokładnie jedna reguła. W naturalny sposób możemy rozważać także schematy niedeterministyczne, definiujące języki drzew - przy czym wówczas będziemy ograniczać się do drzew skończonych. Dla schematu $\mathcal{S}$ język drzew skończonych przez niego generowanych (nie zawierających symbolu ω) oznaczamy przez $L(\mathcal{S})$. Jeżeli wszystkie litery z Σ mają rangę co najwyżej 1, generowane drzewa, składające się z jednej gałęzi, traktujemy jako słowa nad Σ; pomijamy przy tym etykietę liścia (używamy zazwyczaj tylko jednej litery o randze 0, traktowanej jako znak końca słowa).

Przykład 3. (niedeterministyczny):

$(\Sigma, rank) = (\{b, c, d, e\}, \{(b, 1), (c, 1), (d, 1), (e, 0)\})$

$\mathcal{N} = \{S, X, I, C\}$

Zbiór reguły ($\mathcal{R}$):

$S \longrightarrow X\, I\, e$

$X\, f\, x \longrightarrow f\, x$

$X\, f\, x \longrightarrow b\,(X\,(C\, f)\,(d\, x))$

$I\, x \longrightarrow x$

$C\, f\, x \longrightarrow c\,(f\, x)$

Łatwo sprawdzić, że:

$L(\mathcal{S}) = \{b^n c^n d^n \mid n = \{0, 1, \ldots\}\}$.

$$\begin{array}{c} b \\ \left.\begin{array}{c} | \\ \vdots \\ | \end{array}\right]n \\ b \\ | \\ c \\ \left.\begin{array}{c} | \\ \vdots \\ | \end{array}\right]n \\ c \\ | \\ d \\ \left.\begin{array}{c} | \\ \vdots \\ | \end{array}\right]n \\ d \\ | \\ e \end{array}$$

Powiemy, że schemat rekurencyjny jest rzędu r, ozn. $ord(\mathcal{S}) = r$, jeżeli $r = \max\limits_{N \in \mathcal{N}}(ord(N))$. Wiadomo, że gramatyki rzędu 0, rozważane jako generatory słów skończonych, są równoważne automatom skończonym, a rzędu 1 gramatykom bezkontekstowym.