Greibachovej normálny tvar

Peter Kostolányi

7. marca 2017

Definícia a štandardná konštrukcia

V nasledujúcom ukážeme, že ku každej bezkontextovej gramatike existuje (až na ε) ekvivalentná bezkontextová gramatika obsahujúca iba pravidlá s pravou stranou začínajúcou terminálom – o takýchto gramatikách budeme hovoriť, že sú v *Greibachovej normálnom tvare*. V každom kroku odvodenia gramatiky v Greibachovej normálnom tvare sa tak vygeneruje aspoň jeden terminálny symbol výsledného slova. Dôsledkom je, že pre každý bezkontextový jazyk L existuje bezkontextová gramatika, v ktorej každé slovo w z L možno odvodiť na $|w|$ krokov odvodenia. Keďže sa navyše terminály generujú „naľavo“, Greibachovej normálny tvar možno použiť na dôkaz, že ku každému zásobníkovému automatu existuje ekvivalentný, ktorý neobsahuje prechody na prázdne slovo ε.

Definícia 1. Nech $G = (N, T, P, \sigma)$ je bezkontextová gramatika. Hovoríme, že gramatika G je v *Greibachovej normálnom tvare*, ak $P \subseteq N \times T(N \cup T)^*$.

Pomenovanie „normálny tvar“ je odôvodnené nasledujúcou vetou, ktorej dôkaz možno nájsť napríklad v knihe od Hopcrofta a Ullmana [1]: ku každej bezkontextovej gramatike existuje (až na ε) ekvivalentná bezkontextová gramatika v Greibachovej normálnom tvare.

Veta 1. *Nech $G = (N, T, P, \sigma)$ je bezkontextová gramatika. Potom existuje bezkontextová gramatika $G' = (N', T', P', \sigma')$ v Greibachovej normálnom tvare taká, že $L(G') = L(G) - \{\varepsilon\}$.*

Poznámka 1. Niekedy sa v definícii Greibachovej normálneho tvaru povoľuje pravidlo $\sigma \to \varepsilon$ za predpokladu, že σ sa nevyskytuje na pravej strane žiadneho prepisovacieho pravidla. Inými slovami, požaduje sa, aby platilo $P \subseteq (N \times T((N - \{\sigma\}) \cup T)^*) \cup \{\sigma \to \varepsilon\}$. Táto úprava zjavne rieši problém s prázdnym slovom ε a pre túto definíciu Greibachovej normálneho tvaru platí, že ku každej bezkontextovej gramatike G existuje bezkontextová gramatika G' v Greibachovej normálnom tvare taká, že $L(G') = L(G)$.

Vstupom nasledujúceho algoritmu na prevod bezkontextovej gramatiky do Greibachovej normálneho tvaru je bezkontextová gramatika $G = (N, T, P, \sigma)$ v „bezepsilonovom“ normálnom tvare, ktorá neobsahuje reťazové pravidlá („chain rules“). Predpokladajme, že je dané očíslovanie neterminálov gramatiky G; čiže $N = \{\xi_1, \ldots, \xi_n\}$. Dôkaz správnosti algoritmu možno nájsť v [1].

1. Opakuj pre $k = 1, \ldots, n$:
   - 1.1 Opakuj pre $j = 1, \ldots, k-1$:
     - 1.1.1 Opakuj pre všetky pravidlá $\xi_k \to \xi_j w$, kde $w \in (N \cup T)^*$:
       - 1.1.1a Pre každé pravidlo $\xi_j \to u$, kde $u \in (N \cup T)^*$, pridaj pravidlo $\xi_k \to uw$.
       - 1.1.1b Odober pravidlo $\xi_k \to \xi_j w$.
   - 1.2 Pridaj nový neterminál ξ_{k-n}.
   - 1.3 Opakuj pre všetky pravidlá $\xi_k \to \xi_k w$, $w \in (N \cup T)^*$:
     - 1.3a Pridaj pravidlá $\xi_{k-n} \to w\xi_{k-n}$ a $\xi_{k-n} \to w$.
     - 1.3b Odober pravidlo $\xi_k \to \xi_k w$.
   - 1.4 Pre všetky pravidlá $\xi_k \to u$ pridaj pravidlo $\xi_k \to u\xi_{k-n}$.
2. Opakuj pre $k = n-1, \ldots, -(n-1)$:
   - 2.1 Opakuj pre $j = n, \ldots, k+1$:
     - 2.1.1 Opakuj pre všetky pravidlá $\xi_k \to \xi_j w$, kde $w \in (N \cup T)^*$:
       - 2.1.1a Pre každé pravidlo $\xi_j \to u$, kde $u \in (N \cup T)^*$, pridaj pravidlo $\xi_k \to uw$.
       - 2.1.1b Odober pravidlo $\xi_k \to \xi_j w$.

Úloha 1. Nech $G = (N, T, P, \sigma)$ je bezkontextová gramatika s $N = \{\sigma, \alpha\}$, $T = \{a, b\}$ a

$$\begin{aligned} P = \{\sigma &\to \sigma a \mid \alpha\alpha \\ \alpha &\to \sigma\sigma \mid \alpha\alpha \mid b\}. \end{aligned}$$

Štandardnou konštrukciou preveďte gramatiku G do Greibachovej normálneho tvaru.

Riešenie. Po zavedení očíslovania $\sigma =: \xi_1$, $\alpha =: \xi_2$ dostávame gramatiku $G'' = (N'', T, P'', \xi_1)$, kde $N'' = \{\xi_1, \xi_2\}$ a

$$\begin{aligned} P'' = \{\xi_1 &\to \xi_1 a \mid \xi_2\xi_2 \\ \xi_2 &\to \xi_1\xi_1 \mid \xi_2\xi_2 \mid b\}. \end{aligned}$$

Táto gramatika zrejme spĺňa všetky predpoklady na vstup štandardného algoritmu popísaného vyššie. Môžeme preto pristúpiť k simulácii bodu 1.

- Bod 1, $k = 1$. V bode 1.1 zjavne iterujeme cez prázdnu množinu, preto ostane gramatika nezmenená. V bode 1.2 pridáme neterminál ξ_{-1}. V bode 1.3 pridáme pravidlá $\xi_{-1} \to a\xi_{-1}$, $\xi_{-1} \to a$ a odoberieme pravidlo $\xi_1 \to \xi_1 a$. V bode 1.4 následne pridáme pravidlo $\xi_1 \to \xi_2\xi_2\xi_{-1}$. Výsledná sada pravidiel je

$$\begin{aligned} \xi_{-1} &\to a\xi_{-1} \mid a \\ \xi_1 &\to \xi_2\xi_2 \mid \xi_2\xi_2\xi_{-1} \\ \xi_2 &\to \xi_1\xi_1 \mid \xi_2\xi_2 \mid b. \end{aligned}$$

- Bod 1, $k = 2$. V bode 1.1 pridáme pravidlá $\xi_2 \to \xi_2\xi_2\xi_1$, $\xi_2 \to \xi_2\xi_2\xi_{-1}\xi_1$ a odoberieme pravidlo $\xi_2 \to \xi_1\xi_1$. Výsledná sada pravidiel je

$$\begin{aligned} \xi_{-1} &\to a\xi_{-1} \mid a \\ \xi_1 &\to \xi_2\xi_2 \mid \xi_2\xi_2\xi_{-1} \\ \xi_2 &\to \xi_2\xi_2 \mid b \mid \xi_2\xi_2\xi_1 \mid \xi_2\xi_2\xi_{-1}\xi_1. \end{aligned}$$

  V bode 1.2 pridáme neterminál ξ_0. V bode 1.3 pridáme pravidlá $\xi_0 \to \xi_2\xi_0$, $\xi_0 \to \xi_2$, $\xi_0 \to \xi_2\xi_1\xi_0$, $\xi_0 \to \xi_2\xi_1$, $\xi_0 \to \xi_2\xi_{-1}\xi_1\xi_0$, $\xi_0 \to \xi_2\xi_{-1}\xi_1$ a odoberieme pravidlá $\xi_2 \to \xi_2\xi_2$, $\xi_2 \to \xi_2\xi_2\xi_1$ a $\xi_2 \to \xi_2\xi_2\xi_{-1}\xi_1$. V bode 1.4 pridáme pravidlo $\xi_2 \to b\xi_0$. Výsledná sada pravidiel je

$$\begin{aligned} \xi_{-1} &\to a\xi_{-1} \mid a \\ \xi_0 &\to \xi_2\xi_0 \mid \xi_2 \mid \xi_2\xi_1\xi_0 \mid \xi_2\xi_1 \mid \xi_2\xi_{-1}\xi_1\xi_0 \mid \xi_2\xi_{-1}\xi_1 \\ \xi_1 &\to \xi_2\xi_2 \mid \xi_2\xi_2\xi_{-1} \\ \xi_2 &\to b \mid b\xi_0. \end{aligned}$$

Ako si čitateľ istotne všimol, pravidlá z neterminálu ξ_2 sú už požadovaného tvaru. Pristúpme teraz k simulácii bodu 2.

- Bod 2, $k = 1$. Pridáme pravidlá $\xi_1 \to b\xi_2$, $\xi_1 \to b\xi_0\xi_2$, $\xi_1 \to b\xi_2\xi_{-1}$ a $\xi_1 \to b\xi_0\xi_2\xi_{-1}$. Odoberieme pravidlá $\xi_1 \to \xi_2\xi_2$ a $\xi_1 \to \xi_2\xi_2\xi_{-1}$. Výsledná sada pravidiel je

$$\begin{aligned} \xi_{-1} &\to a\xi_{-1} \mid a \\ \xi_0 &\to \xi_2\xi_0 \mid \xi_2 \mid \xi_2\xi_1\xi_0 \mid \xi_2\xi_1 \mid \xi_2\xi_{-1}\xi_1\xi_0 \mid \xi_2\xi_{-1}\xi_1 \\ \xi_1 &\to b\xi_2 \mid b\xi_0\xi_2 \mid b\xi_2\xi_{-1} \mid b\xi_0\xi_2\xi_{-1} \\ \xi_2 &\to b \mid b\xi_0. \end{aligned}$$

- Bod 2, $k = 0$. Pridáme pravidlá $\xi_0 \to b\xi_0$, $\xi_0 \to b\xi_0\xi_0$, $\xi_0 \to b$, $\xi_0 \to b\xi_0$, $\xi_0 \to b\xi_1\xi_0$, $\xi_0 \to b\xi_0\xi_1\xi_0$, $\xi_0 \to b\xi_1$, $\xi_0 \to b\xi_0\xi_1$, $\xi_0 \to b\xi_{-1}\xi_1\xi_0$, $\xi_0 \to b\xi_0\xi_{-1}\xi_1\xi_0$, $\xi_0 \to b\xi_{-1}\xi_1$ a $\xi_0 \to b\xi_0\xi_{-1}\xi_1$. Odoberieme všetkých šesť pôvodných pravidieľ zo ξ_0. Výsledná sada pravidiel (po odstránení duplikátov) je

$$\begin{aligned}
\xi_{-1} &\to a\xi_{-1} \mid a \\
\xi_0 &\to b\xi_0 \mid b\xi_0\xi_0 \mid b \mid b\xi_1\xi_0 \mid b\xi_0\xi_1\xi_0 \mid b\xi_1 \mid b\xi_0\xi_1 \mid b\xi_{-1}\xi_1\xi_0 \mid b\xi_0\xi_{-1}\xi_1\xi_0 \mid b\xi_{-1}\xi_1 \mid b\xi_0\xi_{-1}\xi_1 \\
\xi_1 &\to b\xi_2 \mid b\xi_0\xi_2 \mid b\xi_2\xi_{-1} \mid b\xi_0\xi_2\xi_{-1} \\
\xi_2 &\to b \mid b\xi_0.
\end{aligned}$$

- Bod 2, $k = -1$. Keďže iterujeme cez prázdnu množinu pravidiel, gramatika ostane nezmenená.

Môžeme teda uzavrieť, že gramatika G je ekvivalentná gramatike $G' = (N', T, P', \xi_1)$ v Greibachovej normálnom tvare, kde $N' = \{\xi_{-1}, \xi_0, \xi_1, \xi_2\}$ a

$$\begin{aligned}
P' = \{\xi_{-1} &\to a\xi_{-1} \mid a \\
\xi_0 &\to b\xi_0 \mid b\xi_0\xi_0 \mid b \mid b\xi_1\xi_0 \mid b\xi_0\xi_1\xi_0 \mid b\xi_1 \mid b\xi_0\xi_1 \mid b\xi_{-1}\xi_1\xi_0 \mid b\xi_0\xi_{-1}\xi_1\xi_0 \mid b\xi_{-1}\xi_1 \mid b\xi_0\xi_{-1}\xi_1 \\
\xi_1 &\to b\xi_2 \mid b\xi_0\xi_2 \mid b\xi_2\xi_{-1} \mid b\xi_0\xi_2\xi_{-1} \\
\xi_2 &\to b \mid b\xi_0\}.
\end{aligned}$$ □

Zosilnenia Greibachovej normálneho tvaru

Často sa Greibachovej normálny tvar definuje so silnejšou požiadavkou, aby pre pre množinu prepisovacích pravidiel P platilo $P \subseteq N \times TN^*$. Úprava gramatiky v Greibachovej normálnom tvare (podľa definície 1) na tento silnejší tvar je podobná ako pri prevode do Chomského normálneho tvaru: pre každý terminál c stačí zaviesť nový neterminál ξ_c a pravidlo $\xi_c \to c$ a následne na pravej strane každého pravidla všetky výskytu terminálu c okrem prvého zameniť za ξ_c. Formálnejší popis tejto konštrukcie a dôkaz jej správnosti prenechávame čitateľovi.

Niekedy sa v definícii Greibachovej normálneho tvaru kladie ešte silnejšia požiadavka, aby pre množinu prepisovacích pravidiel platilo $P \subseteq N \times (T \cup TN \cup TNN)$. Dôkaz, že skutočne ide o normálny tvar, je prenechaný ako jedna z úloh na cvičenie č. 4.

„Odepsilonovanie“ zásobníkových automatov

Greibachovej normálny tvar možno využiť na dôkaz, že ku každému zásobníkovému automatu existuje ekvivalentný bez prechodov na prázdne slovo ε. Ľahko možno nahliadnuť, že ak ku gramatike v Greibachovej normálnom tvare pomocou štandardnej konštrukcie zostrojíme ekvivalentný zásobníkový automat, tak ľubovoľná konfigurácia tohto automatu s neterminálom na vrchu zásobníka môže byť nasledovaná iba konfiguráciou, v ktorej je na vrchu zásobníka terminálny symbol. Preto stačí upraviť prechodovú funkciu tak, aby automat tento terminál prečítal zo vstupu už v predchádzajúcom kroku výpočtu.

Veta 2. *Nech $G = (N, T, P, \sigma)$ je bezkontextová gramatika v Greibachovej normálnom tvare. Potom existuje zásobníkový automat $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ taký, že $N(A) = L(G)$ a pre všetky $q \in K$ a $Z \in \Gamma$ platí $\delta(q, \varepsilon, Z) = \emptyset$.*

Dôkaz. Automat A zostrojíme nasledovne: $K = \{q_0\}$, $\Sigma = T$, $\Gamma = N \cup T$, $Z_0 = \sigma$, $F = \emptyset$, pričom pre všetky $\xi \in N$, $c \in T$ a $x \in (N \cup T)^*$ také, že $\xi \to cx \in P$ bude $(q_0, x^R) \in \delta(q_0, c, \xi)$ a pre všetky $c \in T$ bude $(q_0, \varepsilon) \in \delta(q_0, c, c)$. Automat A nemá definované žiadne ďalšie prechody. Dôkaz tvrdenia $N(A) = L(G)$ prenechávame čitateľovi. □

Uvedená konštrukcia sa dá upraviť, aby fungovala aj pre akceptáciu stavom. Tento problém je náplňou jednej z úloh určených na cvičenie č. 4.

Literatúra

[1] Hopcroft, J. E., Ullman, J. D.: *Introduction to Automata Theory, Languages, and Computation.* Reading : Addison-Wesley, 1979. ISBN 0-201-02988-X.