Poznámky k cvičeniu č. 9

Peter Kostolányi

22. novembra 2017

Zásobníkové automaty

Definícia 1. *Zásobníkový automat* (PDA[1]) je sedmica $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, kde K je neprázdna konečná množina stavov, Σ je vstupná abeceda, Γ je abeceda zásobníkových symbolov, $\delta\colon K \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \to 2_{kon}^{K\times\Gamma^*}$ je prechodová funkcia, $q_0 \in K$ je počiatočný stav, $Z_0 \in \Gamma$ je počiatočný zásobníkový symbol a $F \subseteq K$ je množina akceptačných stavov.

Poznámka 1. Symbolom $2_{kon}^{K\times\Gamma^*}$ označujeme množinu *konečných* podmnožín množiny $K \times \Gamma^*$. Obmedzenie sa na konečné podmnožiny je nutné v záujme zachovania konečnosti popisu zásobníkových automatov. Čitateľ sa ľahko presvedčí o tom, že tolerovanie nekonečných podmnožín by malo za následok neúmerný nárast výpočtovej sily zásobníkových automatov, ktoré by takto boli schopné jednoduchým spôsobom akceptovať ľubovoľný jazyk.

Poznámka 2. Pod pojmom *zásobníkový automat* sa zvyčajne chápe *nedeterministický* zásobníkový automat, čo súhlasí s definíciou 1. Deterministický variant zásobníkových automatov sa na tomto predmete zvykne preberať v letnom semestri.

Definícia 2. Nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat. *Konfigurácia* zásobníkového automatu A je trojica (q, w, γ), kde $q \in K$ je stav, $w \in \Sigma^*$ je slovo nad abecedou vstupných symbolov (reprezentujúce nedočítanú časť vstupného slova) a $\gamma \in \Gamma^*$ je slovo nad abecedou zásobníkových symbolov (reprezentujúce obsah zásobníka s dnom naľavo a vrchom napravo).

Poznámka 3. V literatúre sa obsah zásobníka občas zapisuje aj opačne, t.j. s dnom napravo a vrchom naľavo.

Definícia 3. Nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat. *Krok výpočtu* automatu A je binárna relácia $\vdash_A$ na množine konfigurácií automatu A taká, že pre všetky $p, q \in K$, $u, v \in \Sigma^*$ a $\gamma_1, \gamma_2 \in \Gamma^*$ platí $(p, u, \gamma_1) \vdash_A (q, v, \gamma_2)$ práve vtedy, keď existujú $z \in \Sigma \cup \{\varepsilon\}$, $\gamma, \beta \in \Gamma^*$ a $Z \in \Gamma$ tak, že $u = zv$, $\gamma_1 = \gamma Z$, $\gamma_2 = \gamma\beta$ a $(q, \beta) \in \delta(p, z, Z)$. Ak je automat A zrejmý z kontextu, často namiesto $\vdash_A$ píšeme len $\vdash$.

Definícia 4. Nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat. *Jazyk* akceptovaný automatom A *akceptačným stavom* je definovaný ako

$$L(A) = \{w \in \Sigma^* \mid \exists q \in F\ \exists \gamma \in \Gamma^* : (q_0, w, Z_0) \vdash^* (q, \varepsilon, \gamma)\}.$$

Definícia 5. Nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat. *Jazyk* akceptovaný automatom A *prázdnym zásobníkom* je definovaný ako

$$N(A) = \{w \in \Sigma^* \mid \exists q \in K : (q_0, w, Z_0) \vdash^* (q, \varepsilon, \varepsilon)\}.$$

Úloha 1. Nech $L = \{w \in \{a, b\}^* \mid w = w^R\}$. Zostrojte zásobníkový automat A taký, že $N(A) = L$. Správnosť svojej konštrukcie zdôvodnite.

Riešenie. Zostrojíme zásobníkový automat, ktorý prvú polovicu slova uloží počas jej čítania na zásobník a túto informáciu použije pri čítaní druhej polovice slova na overenie, či skutočne ide o reverz prvej polovice. Stred slova nebude automat detegovať, ale v každom kroku prvej fázy bude mať možnosť nedeterministického rozhodnutia o „prepnutí sa“ do druhej fázy. Správnosť tohto rozhodnutia automat overí na konci výpočtu: počet symbolov prečítaných zo vstupu počas druhej fázy sa musí (zhruba) rovnať počtu symbolov na zásobníku pred druhou fázou.

[1] Z angl. *Pushdown automaton.*

Formálna konštrukcia: $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, kde $K = \{q_0, q_1\}$, $\Sigma = \{a, b\}$, $\Gamma = \{Z_0, a, b\}$, množina F je prázdna[2] a prechodová funkcia δ je definovaná nasledovne (na neuvedených vstupoch je výstupom funkcie δ prázdna množina):

$$\begin{aligned}
\delta(q_0, a, Z) &= \{(q_0, Za), (q_1, Z)\}, \quad \forall Z \in \Gamma, \\
\delta(q_0, b, Z) &= \{(q_0, Zb), (q_1, Z)\}, \quad \forall Z \in \Gamma, \\
\delta(q_0, \varepsilon, Z) &= \{(q_1, Z)\}, \quad \forall Z \in \Gamma, \\
\delta(q_1, a, a) &= \{(q_1, \varepsilon)\}, \\
\delta(q_1, b, b) &= \{(q_1, \varepsilon)\}, \\
\delta(q_1, \varepsilon, Z_0) &= \{(q_1, \varepsilon)\}.
\end{aligned}$$

Indukciou (ide tu v podstate o rovnakú metódu invariantov ako pre konečné automaty) by sme ľahko dokázali, že pre slová $u \in \Sigma^*$ a $\gamma \in \Gamma^*$ platí $(q_0, u, Z_0) \vdash^* (q_0, \varepsilon, Z_0\gamma)$ práve vtedy, keď $u = \gamma$. Podobne by sme ľahko dokázali, že pre slová $v \in \Sigma^*$ a $\gamma \in \Gamma^*$ platí $(q_0, v, Z_0) \vdash^* (q_1, \varepsilon, Z_0\gamma)$ práve vtedy, keď existujú slová $x, y \in \Sigma^*$ a $z \in \{a, b, \varepsilon\}$ také, že $v = xyzy^R$ a $\gamma = x$. Nakoniec by sme mohli usúdiť, že pre slovo $w \in \Sigma^*$ platí $(q_0, w, Z_0) \vdash^* (q_1, \varepsilon, \varepsilon)$ práve vtedy, keď $w = yzy^R$ pre nejaké $y \in \Sigma^*$ a $z \in \{a, b, \varepsilon\}$, t.j. ak $w = w^R$. Dôsledkom tejto skutočnosti spoločne s nedosiahnuteľnosťou konfigurácie $(q_0, \varepsilon, \varepsilon)$ je, že $L(A) = L$. □

Ekvivalencia módov akceptácie zásobníkových automatov

Ako bolo dokázané na prednáške, akceptácia stavom a akceptácia prázdnym zásobníkom sú pre zásobníkové automaty ekvivalentné v nasledujúcom zmysle: pre každý zásobníkový automat A existuje zásobníkový automat A' taký, že $N(A') = L(A)$ a zásobníkový automat A'' taký, že $L(A'') = N(A)$.

Veta 1. *Nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat. Potom existuje zásobníkový automat $A' = (K', \Sigma', \Gamma', \delta', q'_0, Z'_0, F')$ taký, že $N(A') = L(A)$.*

Na prednáške bol uvedený dôkaz tejto vety, na základe ktorého možno odvodiť nasledujúci jednoduchý algoritmus. Jeho vstupom je zásobníkový automat A a výstupom je zásobníkový automat A' taký, že $N(A') = L(A)$.

1. Nech Z'_0 je nový zásobníkový symbol (teda $Z'_0 \notin \Gamma$). Ďalej, nech $\Sigma' = \Sigma$, $\Gamma' = \Gamma \cup \{Z'_0\}$, $q'_0 = q_0$ a $F' = \emptyset$.
2. Inicializuj $K' = K$ a $\delta' = \delta$.
3. Pridaj nový prechod $(q_0, Z'_0 Z_0) \in \delta'(q_0, \varepsilon, Z'_0)$.
4. Pridaj do K' nový stav q_z.
5. Pre každé $q \in F$ a $Z \in \Gamma'$ pridaj nový prechod $(q_z, \varepsilon) \in \delta'(q, \varepsilon, Z)$.
6. Pre každé $Z \in \Gamma'$ pridaj nový prechod $(q_z, \varepsilon) \in \delta'(q_z, \varepsilon, Z)$.

Použitie uvedeného jednoduchého postupu na vhodnom príklade zásobníkového automatu prenechávame čitateľovi.

Veta 2. *Nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat. Potom existuje zásobníkový automat $A'' = (K'', \Sigma'', \Gamma'', \delta'', q''_0, Z''_0, F'')$ taký, že $L(A'') = N(A)$.*

Opäť uveďme algoritmus, ktorý možno odvodiť z dôkazu tejto vety. Jeho vstupom je zásobníkový automat A a výstupom je zásobníkový automat A'' taký, že $L(A'') = N(A)$.

[2] Pri konštrukciách zásobníkových automatov zamýšľaných na akceptáciu prázdnou pamäťou sa zvyčajne kladie $F = \emptyset$, keďže množina akceptačných stavov nemá na akceptáciu prázdnou pamäťou žiaden vplyv.

1. Nech Z_0'' je nový zásobníkový symbol (teda $Z_0'' \notin \Gamma$). Ďalej, nech $\Sigma'' = \Sigma$, $\Gamma'' = \Gamma \cup \{Z_0''\}$ a $q_0'' = q_0$.

2. Inicializuj $K'' = K$, $F'' = \emptyset$ a $\delta'' = \delta$.

3. Pridaj nový prechod $(q_0, Z_0''Z_0) \in \delta'(q_0, \varepsilon, Z_0'')$.

4. Pridaj do K'' aj do F'' nový stav q_{fin}.

5. Pre každé $q \in F$ pridaj nový prechod $(q_{fin}, Z_0'') \in \delta''(q, \varepsilon, Z_0'')$.

Konštrukcia zásobníkového automatu k bezkontextovej gramatike

Jednou z kľúčových vlastností zásobníkových automatov je ich ekvivalencia s bezkontextovými gramatikami, ktorá bola dokázaná na prednáške. Trieda jazykov akceptovaných zásobníkovými automatmi je teda práve trieda $\mathscr{L}_{CF}$ všetkých bezkontextových jazykov. V nasledujúcom sa zameriame na simuláciu bezkontextových gramatík pomocou zásobníkových automatov, teda na jednu z inklúzií medzi zodpovedajúcimi triedami jazykov.

Veta 3. *Nech $G = (N, T, P, \sigma)$ je bezkontextová gramatika. Potom existuje zásobníkový automat $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ taký, že $N(A) = L(G)$.*

Automat A ekvivalentný gramatike G pracuje tak, že na zásobníku simuluje ľavé krajné odvodenie gramatiky G, pričom obsah zásobníka čítaný „zhora nadol“ (teda reverz obsahu zásobníka pri zvyčajnom zápise v konfiguráciách) reprezentuje vždy určitý sufix vetnej formy (konkrétne ten, ktorý sa začína prvým výskytom neterminálu; terminálne symboly pred ním sa už počas odvodenia ďalej nemenia).

Na začiatku výpočtu je na zásobníku počiatočný neterminál σ. Ak je na vrchu zásobníka neterminál, prepíše sa pomocou niektorého z pravidiel a zo vstupu sa neprečíta nič. To zodpovedá simulácii jedného kroku odvodenia gramatiky G. Ak je na vrchu zásobníka terminál, automat nemôže priamo simulovať krok odvodenia gramatiky G. Daný terminálny symbol sa už však počas odvodenia meniť nebude, a preto ho môže automat konfrontovať so svojím vstupom. Ak je symbol na vstupe rovnaký, automat ho prečíta a zmaže vrch zásobníka. V opačnom prípade sa zasekne. Vyprázdnenie zásobníka zodpovedá vygenerovaniu terminálneho slova gramatikou G. V prípade, že toto slovo zodpovedá kompletnému vstupu, automat akceptuje prázdnym zásobníkom. Na konštrukciu takéhoto automatu zjavne stačí jediný stav.

Algoritmus konštrukcie zásobníkového automatu $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ k bezkontextovej gramatike $G = (N, T, P, \sigma)$ možno zhrnúť nasledovne:

1. Nech $K = \{q_0\}$, $\Sigma = T$, $\Gamma = N \cup T$, $Z_0 = \sigma$ a $F = \emptyset$.

2. Pre každé pravidlo $\xi \to x$, kde $\xi \in N$ a $x \in (N \cup T)^*$, nech $(q_0, x^R) \in \delta(q_0, \varepsilon, \xi)$.

3. Pre každé $c \in T$, nech $(q_0, \varepsilon) \in \delta(q_0, c, c)$.

4. Automat A nemá definované žiadne ďalšie prechody.

Úloha 2. Nech $G = (N, T, P, \sigma)$ je bezkontextová gramatika s $N = \{\sigma, \alpha, \beta, \gamma\}$, $T = \{a, b\}$ a

$$\begin{aligned} P = \{\sigma &\to a\alpha\beta \mid \alpha\beta \mid \varepsilon \\ \alpha &\to \alpha b\beta \mid aa \\ \beta &\to b\beta\gamma \mid ab \\ \gamma &\to a\gamma \mid aab \mid \varepsilon\}. \end{aligned}$$

Štandardnou konštrukciou zostrojte zásobníkový automat A taký, že $N(A) = L(G)$.

Riešenie. Definujeme $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, kde $K = \{q_0\}$, $\Sigma = \{a, b\}$, $\Gamma = \{a, b, \sigma, \alpha, \beta, \gamma\}$, $Z_0 = \sigma$, $F = \emptyset$, a kde prechodová funkcia δ je daná nasledovne:

$$\begin{aligned}
\delta(q_0, \varepsilon, \sigma) &= \{(q_0, \beta\alpha a), (q_0, \beta\alpha), (q_0, \varepsilon)\}, \\
\delta(q_0, \varepsilon, \alpha) &= \{(q_0, \beta b\alpha), (q_0, aa)\}, \\
\delta(q_0, \varepsilon, \beta) &= \{(q_0, \gamma\beta b), (q_0, ba)\}, \\
\delta(q_0, \varepsilon, \gamma) &= \{(q_0, \gamma a), (q_0, baa), (q_0, \varepsilon)\}, \\
\delta(q_0, a, a) &= \{(q_0, \varepsilon)\}, \\
\delta(q_0, b, b) &= \{(q_0, \varepsilon)\}.
\end{aligned}$$

Pre automat A platí $N(A) = L(G)$. □

Konštrukcia bezkontextovej gramatiky k zásobníkovému automatu

Vyššie sme načali problematiku ekvivalencie zásobníkových automatov s bezkontextovými gramatikami a rozobrali sme konštrukciu zásobníkového automatu ekvivalentného danej bezkontextovej gramatike. V nasledujúcom sa zameriame na opačnú konštrukciu, ktorej vstupom je zásobníkový automat A a výstupom je bezkontextová gramatika $G = (N, T, P, \sigma)$ taká, že $L(G) = N(A)$.

Veta 4. *Nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat. Potom existuje bezkontextová gramatika $G = (N, T, P, \sigma)$ taká, že $L(G) = N(A)$.*

Gramatika G, ekvivalentná automatu A, obsahuje pre všetky dvojice stavov $p, q \in K$ a všetky zásobníkové symboly $Z \in \Gamma$ neterminál $[p, Z, q]$. Množina pravidiel bude neskôr skonštruovaná tak, aby pre terminálne slovo $w \in T^*$ platilo $[p, Z, q] \Rightarrow^* w$ práve vtedy, keď $(p, w, Z) \vdash^* (q, \varepsilon, \varepsilon)$. Tento zápis vyjadruje, že z neterminálu $[p, Z, q]$ sa dajú vygenerovať práve všetky terminálne slová w také, že existuje výpočet automatu A na slove w začínajúci v konfigurácii so stavom p a so symbolom Z na vrchu zásobníka a končiaci v konfigurácii so stavom q, v ktorom sa výška zásobníka automatu A po prečítaní slova w *po prvý raz* dostane o úroveň nižšie v porovnaní s jeho začiatkom.

Keďže automat A akceptuje prázdnou pamäťou, je z uvedeného zrejmé, že gramatika G musí generovať práve všetky slová generované z neterminálov $[q_0, Z_0, q]$ pre $q \in K$. Inými slovami, pre každé $q \in K$ je v množine P pravidlo $\sigma \to [q_0, Z_0, q]$. Zostáva definovať pravidlá zabezpečujúce „správanie“ neterminálov $[p, Z, q]$, opísané v predchádzajúcom odstavci. Zrejme, ak pre nejaké $p, q \in K$, $z \in \Sigma \cup \{\varepsilon\}$ a $Z \in \Gamma$ platí $(q, \varepsilon) \in \delta(p, z, Z)$, musí podľa predchádzajúceho odstavca platiť $[p, Z, q] \Rightarrow^* z$. Preto bude množina P obsahovať pravidlo $[p, Z, q] \to z$. Ak pre nejaké $p, r \in K$, $z \in \Sigma \cup \{\varepsilon\}$ a $Z, Z_1, \ldots, Z_k \in \Gamma$ platí $(r, Z_1 \ldots Z_k) \in \delta(p, z, Z)$, dôjde k zníženiu výšky zásobníka pod pôvodnú úroveň až potom, čo postupne príde k zníženiu pod úroveň novo pridaného Z_k, pod úroveň novo pridaného Z_{k-1}, atď. až pod úroveň novo pridaného Z_1. Stavy, v ktorých k týmto zníženiam úrovne dôjde, môžu byť ľubovoľné. Preto pre všetky stavy $q, q_1, \ldots, q_{k-1} \in K$ bude množina P obsahovať pravidlo $[p, Z, q] \to z[r, Z_k, q_1][q_1, Z_{k-1}, q_2] \ldots [q_{k-1}, Z_1, q]$.

Z týchto úvah (ktoré boli formalizované na prednáške) možno odvodiť nasledujúci postup konštrukcie, ktorého vstupom je zásobníkový automat $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ a výstupom je bezkontextová gramatika $G = (N, T, P, \sigma)$ taká, že $L(G) = N(A)$.

1. Nech $N = \{[p, Z, q] \mid p, q \in K;\ Z \in \Gamma\} \cup \{\sigma\}$ a $T = \Sigma$.

2. Pre každé $q \in K$ obsahuje P pravidlo $\sigma \to [q_0, Z_0, q]$.

3. Pre každé $p, q \in K$, $z \in \Sigma \cup \{\varepsilon\}$ a $Z \in \Gamma$ také, že $(q, \varepsilon) \in \delta(p, z, Z)$, obsahuje P pravidlo $[p, Z, q] \to z$.

4. Pre každé $p, r \in K$, $z \in \Sigma \cup \{\varepsilon\}$ a $Z, Z_1, \ldots, Z_k \in \Gamma$ také, že $(r, Z_1 \ldots Z_k) \in \delta(p, z, Z)$, obsahuje P pre všetky $q, q_1, \ldots, q_{k-1} \in K$ pravidlo $[p, Z, q] \to z[r, Z_k, q_1] \ldots [q_{k-1}, Z_1, q]$.

5. Množina P neobsahuje žiadne iné pravidlá.

Úloha 3. Nech $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je zásobníkový automat, kde $K = \{q_0, q_1\}$, $\Sigma = \{a, b\}$, $\Gamma = \{Z_0, c\}$, $F = \emptyset$ a

$$\begin{aligned}\delta(q_0, a, Z_0) &\ni (q_1, \varepsilon)\\ \delta(q_0, \varepsilon, c) &\ni (q_0, Z_0)\\ \delta(q_1, b, Z_0) &\ni (q_1, Z_0cc).\end{aligned}$$

Nech automat A neobsahuje žiadne ďalšie prechody.[3] Štandardnou konštrukciou zostrojte bezkontextovú gramatiku G takú, že $L(G) = N(A)$.

Riešenie. Gramatika $G = (N, T, P, \sigma)$ bude daná nasledovne:

$$N = \{\sigma, [q_0, Z_0, q_0], [q_0, c, q_0], [q_0, Z_0, q_1], [q_0, c, q_1], [q_1, Z_0, q_0], [q_1, c, q_0], [q_1, Z_0, q_1], [q_1, c, q_1]\},$$

$T = \{a, b\}$ a množina P obsahuje nasledujúce pravidlá:

$$\begin{aligned}\sigma &\to [q_0, Z_0, q_0] \mid [q_0, Z_0, q_1],\\ [q_0, Z_0, q_1] &\to a,\\ [q_0, c, q_0] &\to [q_0, Z_0, q_0],\\ [q_0, c, q_1] &\to [q_0, Z_0, q_1],\\ [q_1, Z_0, q_0] &\to b[q_1, c, q_0][q_0, c, q_0][q_0, Z_0, q_0] \mid b[q_1, c, q_0][q_0, c, q_1][q_1, Z_0, q_0] \mid\\ &\quad \mid b[q_1, c, q_1][q_1, c, q_0][q_0, Z_0, q_0] \mid b[q_1, c, q_1][q_1, c, q_1][q_1, Z_0, q_0],\\ [q_1, Z_0, q_1] &\to b[q_1, c, q_0][q_0, c, q_0][q_0, Z_0, q_1] \mid b[q_1, c, q_0][q_0, c, q_1][q_1, Z_0, q_1] \mid\\ &\quad \mid b[q_1, c, q_1][q_1, c, q_0][q_0, Z_0, q_1] \mid b[q_1, c, q_1][q_1, c, q_1][q_1, Z_0, q_1].\end{aligned}$$

Pravidlo v druhom riadku pritom zodpovedá prechodu $\delta(q_0, a, Z_0) \ni (q_1, \varepsilon)$, pravidlá v treťom a štvrtom riadku zodpovedajú prechodu $\delta(q_0, \varepsilon, c) \ni (q_0, Z_0)$ a pravidlá v piatom a šiestom riadku zodpovedajú prechodu $\delta(q_1, b, Z_0) \ni (q_1, Z_0cc)$. □

Pumpovacia lema pre bezkontextové jazyky

Nech L je *ľubovoľný* bezkontextový jazyk. Potom *existuje* bezkontextová gramatika $G = (N, T, P, \sigma)$ taká, že $L(G) = L$. Nech najdlhšia pravá strana pravidla z P má dĺžku k. Potom v každom kroku odvodenia možno jeden neterminál prepísať na najviac k symbolov. Preto každý strom odvodenia slova dĺžky aspoň k^n má hĺbku najmenej n a obrátene, strom odvodenia s hĺbkou najviac n zodpovedá slovu dĺžky maximálne k^n. Definujme $p = q := k^{|N|+1}$.

Nech $w \in L$ je *ľubovoľné* slovo také, že $|w| \geq p$. Vezmime strom najkratšieho[4] odvodenia slova w v gramatike G. Ten má hĺbku aspoň $|N| + 1$. Najdlhšia cesta z koreňa do niektorého z listov preto obsahuje aspoň $|N| + 2$ uzlov, pričom v liste je terminál alebo prázdne slovo a vo vnútorných uzloch, ktorých je aspoň $|N| + 1$, sú neterminály. Na takejto ceste sa preto musí zopakovať niektorý neterminál ξ. Ak je cesta dlhšia ako $|N| + 2$, predpokladajme, že opakovaný výskyt neterminálu ξ je v rámci najnižších $|N| + 2$ úrovní. Situácia je znázornená na obrázku 1.

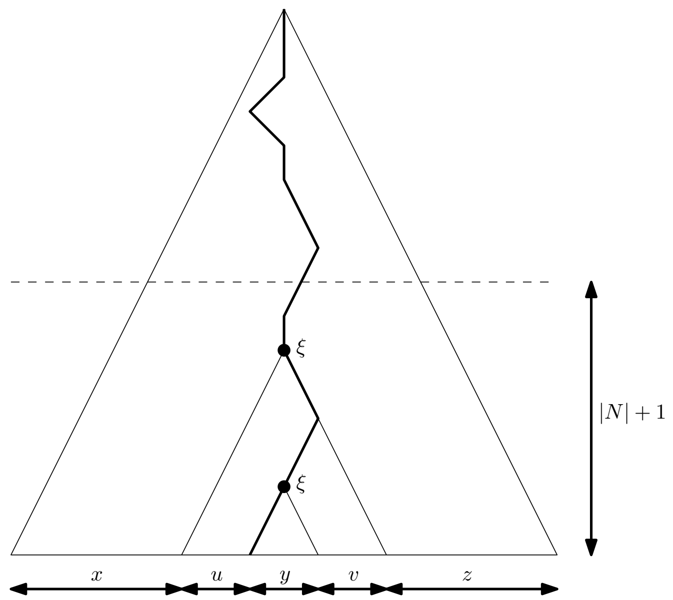
Odvodenie slova w v gramatike G teda možno s odkazom na obrázok 1 zapísať nasledovne:

$$\sigma \Rightarrow^* x\xi z \Rightarrow^* xu\xi vz \Rightarrow^* xuyvz = w.$$

Pre slovo w teda *existujú* slová x, u, y, v, z také, že $w = xuyvz$. Keďže podslovo uyv je vygenerované v podstrome hĺbky $|N| + 1$, platí $|uyv| \leq k^{|N|+1} = q$. Ak by boli obidve slová u, v prázdne, bolo by možné podvýpočet $x\xi z \Rightarrow^* xu\xi vz$ vynechať bez toho, aby sa zmenilo vygenerované slovo, čo je spor s minimalitou odvodenia slova w. Preto $|uv| \geq 1$. A nakoniec, keďže podvýpočet $x\xi z \Rightarrow^* xu\xi vz$ možno vynechať alebo niekoľkokrát zopakovať, pre každé $i \in \mathbb{N}$ musí aj slovo xu^iyv^iz patriť do jazyka generovaného gramatikou G. Dokázali sme teda nasledujúce tvrdenie (pumpovaciu lemu pre bezkontextové jazyky):

[3] V tejto úlohe nejde o zmysluplný automat, ale o čo najmenej bolestnú demonštráciu štandardnej konštrukcie.

[4] Namiesto termínov „najkratší“ a „najdlhší“, používaných v tomto texte, by asi bolo presnejšie používať termíny „minimálny“ a „maximálny“.



Obr. 1: Schematické znázornenie stromu odvodenia v gramatike G. Ak má slovo dĺžku aspoň p, jeho strom odvodenia je hĺbky najmenej $|N|+1$. Najdlhšia cesta z koreňa do listu je znázornená hrubou čiarou. V rámci posledných $|N|+2$ uzlov (čiže v podstrome hĺbky $|N|+1$) sa na tejto ceste musí zopakovať niektorý neterminál ξ. Tieto jeho dva výskyty určujú faktorizáciu $w = xuyvz$ vygenerovaného slova.

Veta 5. *Nech $L \in \mathscr{L}_{CF}$ je bezkontextový jazyk. Potom existujú $p, q \in \mathbb{N}$ také, že pre všetky $w \in L$, $|w| \geq p$, existujú slová $x, u, y, v, z \in \Sigma_L^*$ také, že platí:*

(i) $w = xuyvz$,

(ii) $|uyv| \leq q$,

(iii) $|uv| \geq 1$,

(iv) $\forall i \in \mathbb{N} : xu^iyv^iz \in L$.

Dokazovanie tvrdení typu $L \notin \mathscr{L}_{CF}$

Pumpovacia lema pre bezkontextové jazyky je užitočným nástrojom najmä v prípadoch, keď je potrebné dokázať, že nejaký jazyk L nie je bezkontextový. Jej použitie je podobné ako pri pumpovacej leme pre regulárne jazyky a ilustrujeme ho na nasledujúcej ukážkovej úlohe.

Úloha 4. Dokážte, že jazyk $L = \{wcw \mid w \in \{a,b\}^*\}$ nie je bezkontextový.

Riešenie. Za účelom sporu predpokladajme, že $L \in \mathscr{L}_{CF}$. Nech p, q sú konštanty zodpovedajúce jazyku L podľa pumpovacej lemy. Uvažujme slovo $w = a^{p+q}b^{p+q}ca^{p+q}b^{p+q} \in L$. Z pumpovacej lemy potom vyplýva, že existujú slová x, u, y, v, z také, že platí (i) až (iv). Zrejme musí platiť $\#_c(u) = \#_c(v) = 0$. Keďže $|uyv| \leq q$, musí nastať jedna z nasledujúcich situácií:

a) Platí $|x|+|u|+|y|+|v| \leq 2(p+q)$. Potom zrejme $xu^2yv^2z = w_1cw_2$ pre nejaké $w_1, w_2 \in \{a,b\}^*$ také, že $|w_1| > |w_2|$. Preto $xu^2yv^2z \notin L$.

b) Platí $|u|+|y|+|v|+|z| \leq 2(p+q)$. Potom zrejme $xu^2yv^2z = w_1cw_2$ pre nejaké $w_1, w_2 \in \{a,b\}^*$ také, že $|w_1| < |w_2|$. Preto $xu^2yv^2z \notin L$.

c) Pre nejaké čísla $k, l, m, n \in \mathbb{N}$ platí $x = a^{p+q}b^k$, $u = b^l$, $y = b^{p+q-k-l}ca^m$, $v = a^n$ a $z = a^{p+q-m-n}b^{p+q}$, pričom aspoň jedno z čísel l a n je kladné. Potom zrejme

$$xu^2yv^2z = a^{p+q}b^{p+q+l}ca^{p+q+n}b^{p+q} \notin L. \qquad \square$$