Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Neuniformné modely výpočtov

Bakalárska práca

2021

Ján Priner

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Neuniformné modely výpočtov

Bakalárska práca

Študijný program:	Informatika
Študijný odbor:	Informatika
Školiace pracovisko:	Katedra informatiky
Školiteľ:	prof. RNDr. Rastislav Kráľovič, PhD.

Bratislava, 2021
Ján Priner

24492450

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Ján Priner
Študijný program: informatika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Neuniformné modely výpočtov
Non-uniform models of computation

Anotácia: V sérii nedávnych článkov sa skúmali vlastnosti neuniformných konečných automatov v modeli s poradnou páskou. V tomto modeli sa uvažuje dvojica konečný
automat a poradná funkcia, čo je ľubovoľná funkcia z prirodzených čísel do reťazcov. Neuniformý automat má dodatočnú pásku, na ktorej má prístupnú hodnotu
poradnej funkcie pre dĺžku vstupného slova. Ukázalo sa, že v neuniformnej verzii je rozdiel vo výrazovej sile foriem automatov, ktorých uniformné verzie sú ekvivalentné (napr. nedeterministické jednosmerné automaty vedia rozpoznávať všetky
jazyky, ale deterministické nie). Cieľom práce je preskúmať, či sa tieto vzťahy prenášajú na iné výpočtové modely, napr. na zásobníkové automaty. Alternatívne
sa dá preskúmať súvislosť s podobným modelom definovaným Rusinsom Freivaldsom

Vedúci: prof. RNDr. Rastislav Kráľovič, PhD.
Katedra: FMFI.KI - Katedra informatiky
Vedúci katedry: prof. RNDr. Martin Škoviera, PhD.

Dátum zadania: 04.11.2021

Dátum schválenia: 04.11.2021 doc. RNDr. Daniel Olejár, PhD.
garant študijného programu

....
študent vedúci práce

Abstrakt

Táto práca sa venuje modelu zásobníkových automatov s pomocnou informáciou na dodatočnej páske závislej od dĺžky vstupného slova. Porovnáme silu tohto modelu s modelom konečných automatov s poradnou páskou zavedeným Küçükom et al. Ukážeme existenciu hierarchií v závislosti od dĺžky pomocnej informácie. Preskúmame uzáverové vlastnosti tohto modelu. Ukážeme ako sa vlastnosti modelu líšia medzi deterministickým a nedeterministickým variantom nášho modelu.

Kľúčové slová: zásobníkový automat, konečný automat, pomocná informácia, poradná páska

Abstract

We consider a model of push-down automata with advice string on a separate tape which depends only on the length of the input word. We compare the power of this model with the model of finite-state automata with an advice tape introduced by Küçük et al. We show the existence of hierarchies of advice lengths. We explore the closure properties of this model. Also, we show how the properties of the model differ between the deterministic and nondeterministic variants of our model.

Keywords: push-down automata, finite-state automata, advised computation, advice tape

Obsah

Úvod

Uniformita je vlastnosť výpočtových modelov, kde sa na spracovanie vstupov ľubovoľnej dĺžky používa to isté, konečne popísateľné zariadenie. Niektoré výpočtové modely sú prirodzene uniformné, napríklad konečné automaty, zatiaľčo iné, napríklad logické obvody sú inherentne neuniformné. V minulosti boli skúmané spôsoby, ako neuniformné modely spraviť uniformnými (napr. logické obvody vygenerované Turingovym strojom), a ako uniformné modely spraviť neuniformnými (pomocou pomocnej informácie závislej od dĺžky vstupu).

Karp a Lipton[4] zaviedli model Turingovych strojov ktoré dostanú pomocnú informáciu (radu) prefixovanú pred vstupné slovo. Pre všetky slová jednej dĺžky n dostane Turingov stroj rovnakú radu $h(n)$.

Damm a Holzer[1] rozšírili koncept rady prefixovanej pred vstupné slovo aj na slabšie modely. Ukázali separáciu medzi triedami jazykov s radou konštantnej dĺžky a polynomiálnej dĺžky, a medzi konečnými, zásobníkovými a lineárne ohraničenými automatmi.

Modely v ktorých je rada prefixovaná pred vstupom sú značne limitované. Konečné automaty s prefixovanou radou nedokážu zúžitkovať radu väčšej než konštantnej dĺžky[1]. Preto sa zaviedol model v ktorom je rada na dodatočnej páske. Küçük a Say[5] skúmali takýto model deterministických konečných automatov. Ďuriš et al.[2][3] nadviazali na tento výskum. Venovali sa nedeterministických konečných automatov. Ukázali tiež, že deterministické konečné automaty nedokážu využiť dlhšiu ako polynomiálnu radu.

V našej práci budeme skúmať model neuniformných zásobníkových automatov s pomocnou informáciou na dodatočnej páske. Porovnáme jeho silu s konečnými automatmi predstavenými v [5]. Ukážeme existenciu hierarchií tried jazykov akceptovaných zásobníkovými automatmi s rôznou dĺžkou pomocnej informácie. Pozrieme sa na uzáverové vlastnosti týchto tried. V práci sa venujeme deterministickým aj nedeterministickým zásobníkovým automatom, ukážeme ako sa ich vlastnosti navzájom líšia.

Kapitola 1

Model a základné definície

Náš model je vcelku priamočiarym rozšírením modelu aký používa Küçük [5] na zásobníkové automaty. Máme dve pásky, jednu vstupnú, a jednu s pomocnou informáciou. Pásky sa dajú iba čítať. Vstupná páska aj poradná páska sú ukončené špeciálnym koncovým znakom \$. Zásobníkový automat je osmica $A = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F, \Delta)$. Význam K, Σ, Γ, q_0, Z_0, F je rovnaký ako pri bežných zásobníkových automatoch. Δ je abeceda poradnej pásky. $\delta : K \times (\Sigma \cup \{\$\}) \times \Gamma \times (\Delta \cup \{\$\}) \to 2_{\text{kon}}^{K \times \Gamma^* \times \{0,1\} \times \{0,1\}}$ je prechodová funkcia. Posledné dve zložky určujú, či sa automat posunie vpred na vstupnej a poradnej páske. Ak ide o deterministický automat, navyše vyžadujeme, že všetky hodnoty prechodovej funkcie δ sú najviac jednoprvkové množiny.

Poradná funkcia je funkcia z prirodzených čísel $\mathbb{N}$ na slová nad poradnou abecedou Δ. Značíme ju h. Počas výpočtu slova w sa na poradnej páske bude nachádzať slovo $h(|w|)$. Ak sa automat A pri výpočte v ktorom je na vstupnej páske w, a na poradnej páske $h(|w|)$ dostane do konfigurácie, kde sa hlava na vstupnej páske nachádza na koncovom znaku \$ a automat je v akceptačnom stave, tak povieme, že A s poradnou funkciou h akceptuje slovo w.

Jazyk akceptovaný automatom A s poradnou funkciou h označujeme $L(A, h)$. Motiváciou pre takéto značenie je, že bude užitočné sa pozerať na jazyky ktoré akceptuje ten istý automat s rôznymi poradnými funkciami.

Veta 1.1 (normálny tvar)**.** *Pre každý nedeterministický zásobníkový automat A existuje ekvivalentný nedeterministický zásobníkový automat A' taký, že $L(A', h) = L(A, h)$ pre každé h, a A' naviac spĺňa tieto podmienky:*

1. *V každom kroku automat A' buď pridá jeden znak na zásobník, alebo odoberie vrchný znak zo zásobníka, alebo nechá zásobník rovnaký. Teda ak $(p, \gamma, v_1, v_2) \in \delta_{A'}(q, c, Z, d)$ tak $\gamma \in \{\varepsilon\} \cup \{Zx \mid x \in \Gamma \cup \{\varepsilon\}\}$*

2. *Automat A' sa presunie do akceptačného stavu iba ak má prázdny zásobník.*

Dôkaz. Ak má automat A prechod $(p, Z_1Z_2 \dots Z_k, v_1, v_2) \in \delta_A(q, c, Z, d)$, tak automat A' tento prechod spraví v $k+1$ krokoch. Použije na to novozavedené stavy $q_1, q_2, \dots, q_k$.

$$
\begin{aligned}
(q_1, \varepsilon, 0, 0) &\in \delta_{A'}(q, c, Z, d) \\
(\forall i \in \{1, 2, \dots, k-1\})(\forall Z' \in \Gamma)\ \delta_{A'}(q_i, c, Z', d) &= \{(q_{i+1}, Z'Z_i, 0, 0)\} \\
\delta_{A'}(q_k, c, Z_{k-1}, d) &= \{(p, Z_{k-1}Z_k, v_1, v_2)\}
\end{aligned}
$$

Toto zaručí že automat A' bude spĺňať prvú podmienku. Na to aby automat A' spĺňal aj druhú podmienku zavedieme dva nové stavy q_e a q_f. q_f je jediný akceptačný stav automatu A'. Ak sa čítacia hlava nachádza na koncovom znaku \$ a automat sa nachádza v stave ktorý bol akceptačný pre automat A, tak prejde do stavu q_e. Potom bude automat vyprázdňovať zásobník až dokým nenarazí na spodný znak Z_0. Posledným krokom vyberie Z_0 zo zásobníku a prejde do stavu q_f.

$$
\begin{aligned}
(\forall q \in F_A)(\forall Z \in \Gamma)(\forall d \in \Delta \cup \{\$\})\ \delta_{A'}(q, \$, Z, d) &= \{(q_e, Z, 0, 0)\} \\
(\forall Z \in \Gamma)(\forall d \in \Delta \cup \{\$\})\ \delta_{A'}(q_e, \$, Z, d) &= \{(q_e, \varepsilon, 0, 0)\} \\
(\forall d \in \Delta \cup \{\$\})\ \delta_{A'}(q_e, \$, Z_0, d) &= \{(q_f, \varepsilon, 0, 0)\}
\end{aligned}
$$

Ak bol automat A deterministický tak takto skonštruovaný automat A' je tiež deterministický. □

Kapitola 2

Hierarchie podľa dĺžky rady

V tejto kapitole sa budeme venovať vplyvu veľkosti rady na silu zásobníkových automatov. Pozrieme sa tiež na vzťahy medzi triedami konečných a zásobníkových automatov s radou.

Triedu jazykov akceptovaných zásobníkovým automatom, ktorý používa poradnú pásku dlhú $O(f(n))$ pre slová dlhé n budeme značiť $\mathsf{CFL}/f(n)$, ak ide o nedeterministické automaty a $\mathsf{DCFL}/f(n)$, ak ide o deterministické automaty. Pre deterministické konečné automaty budeme používať označenie $\mathsf{SPACE}(1)/f(n)$, aby sme boli konzistentní s [5].

Namiesto $f(n)$ budeme tiež používať *poly* a *exp*, keď budeme hovoriť vo všeobecnosti o radách polynomiálnej, a exponenciálnej dĺžky ($\mathsf{CFL}/\mathit{poly}$ sú nedeterministické zásobníkové automaty s polynomiálnou radou).

Triedu jazykov akceptovaných nedeterministickým zásobníkovým automatom s radou konštantnej dĺžky k, s binárnou poradnou abecedou budeme značiť CFL/k (používame k alebo číslo keď ide o konštantu). $\mathsf{CFL}/0 = \mathsf{CFL}$. Rovnakým spôsobom budeme chápať aj DCFL/k a $\mathsf{SPACE}(1)/k$.

Veta 2.1. $\mathsf{DCFL}/\mathit{exp} = \mathsf{CFL}/\mathit{exp} = \mathsf{ALL}$

Dôkaz. Trieda $\mathsf{CFL}/\mathit{exp}$ je nadmnožinou $\mathsf{DCFL}/\mathit{exp}$. Teda nám stačí dokázať rovnosť $\mathsf{DCFL}/\mathit{exp} = \mathsf{ALL}$. Inak povedané, stačí nám dokázať, že každý jazyk L je možné akceptovať deterministickým zásobníkovým automatom s exponenciálne dlhou radou.

Nech L je jazyk nad abecedou $\Sigma = \{c_1, c_2, \ldots, c_k\}$. Skonštruujeme zásobníkový automat akceptujúci L.

Rada pre slová dlhé n bude k-árny strom hĺbky n. Listy budú znak 0 alebo 1 a podstromy budú ohraničené zátvorkami. Keď automat prečíta znak c_i zo vstupnej pásky, vnorí sa do i-teho dieťaťa v strome na poradnej páske (preskočí $i-1$ podstromov, zásobník používa k tomu, aby si počítal zátvorky). Keď dôjde na koniec vstupu, bude sa hlava na poradnej páske nachádzať na liste stromu. Takto každý list korešponduje k inému slovu w. Rada bude mať znak 1 v listoch ktoré korešpondujú k slovám patriacim do jazyka L, a 0 v listoch korešpondujúcim k slovám ktoré nepatria do L. Automat akceptuje, ak sa po prečítaní celého vstupu nachádza na liste so znakom 1. □

Nakoľko automaty s exponenciálnou radou dokážu akceptovať každý jazyk, nebudeme sa nimi ďalej zaoberať.

Veta 2.2. *Pre každé $k \in \mathbb{N}$, existuje jazyk L_k taký, že L_k patrí do $\mathsf{SPACE}(1)/k+1$, ale nepatrí do CFL/k.*

Dôkaz. Dôkaz bude konštruktívny, popíšeme jazyk L_k a ukážeme, že spĺňa tvrdenie vety. Jazyk L_k bude podmnožinou $\{b^l a^{n-l} \mid n, l \in \mathbb{N},\ l < k+1\}$.

Máme kódovanie nedeterministických zásobníkových automatov na prirodzené čísla (poradnú funkciu nekódujúc). Nezáleží na tom aké, predpokladáme nejaké fixné kódovanie. Automat, ktorý sa kóduje na číslo m budeme značiť A_m. Automat A_m spolu s poradnou funkciou h nám určuje jazyk $L(A_m, h)$, ktorý automat akceptuje.

Pre kód m množina $L(A_m, h) \cap \{a, b\}^m$ obsahuje iba slová dĺžky m, a teda závisí iba od hodnoty $h(m)$. Pre poradné funkcie $h : \mathbb{N} \to \{0,1\}^k$ je len 2^k možností ako môže vyzerať množina $L(A_m, h) \cap \{a, b\}^m$.

Nech $L_{k,m}$ je taká podmnožina množiny $\{b^l a^{m-l} \mid 0 \leq l < k+1\}$ (spomedzi 2^{k+1} možných podmnožín) ktorá sa nezhoduje s $L(A_m, h) \cap \{a, b\}^m$ pre žiadne $h : \mathbb{N} \to \{0,1\}^k$.

Jazyk L_k zostrojíme takto:

$$L_k = \bigcup_{m=0}^{\infty} L_{k,m}$$

Teraz ukážeme, že L_k sa líši od každého jazyka z CFL/k. Nech $L \in \mathsf{CFL}/k$. Musí existovať automat A, a poradná funkcia $h : \mathbb{N} \to \{0,1\}^k$, taká že $L = L(A, h)$. Nech m je kód automatu A. $L_k \cap \{a, b\}^m = L_{k,m}$. Z konštrukcie je zjavné, že $L \cap \{a, b\}^m \neq L_{k,m}$. Keďže $L \cap \{a, b\}^m$ sa nezhoduje s $L_k \cap \{a, b\}^m$, jazyky L a L_k musia byť rôzne.

Konečný automat akceptujúci L_k bude vyzerať nasledovne: Pre slová dlhé n dostane automat radu $c_0 c_1 \dots c_{k-1}$, kde c_i je 1, ak slovo $b^i a^{n-i}$ patrí do jazyka L_k, a c_i je 0, ak slovo $b^i a^{n-i}$ nepatrí do jazyka L_k.

Automat sa na poradnej páske posunie toľkokrát, koľko je znakov b na začiatku vstupného slova. Potom čo narazí na prvé a, skontroluje či všetky ďalšie znaky sú a. Potom akceptuje, ak na poradnej páske prečíta znak 1. Rada nám presne popisuje, aká podmnožina slov tvaru $\{b^l a^{n-l} \mid n, l \in \mathbb{N},\ l < k+1\}$ patrí do jazyka. □

Dôsledok 2.2.1. *Pre každé $k \in \mathbb{N}$, $\mathsf{CFL}/k \subsetneq \mathsf{CFL}/k+1$*

Dôsledok 2.2.2. *Pre každé $k \in \mathbb{N}$, $\mathsf{DCFL}/k \subsetneq \mathsf{DCFL}/k+1$*

Pre jazyk L nad abecedou Σ a čísla $n, k \in \mathbb{N}$ definujeme reláciu $\equiv_{L,n,k}$ nad množinou Σ^k takto: $x \equiv_{L,n,k} y$ práve vtedy, keď pre všetky z dĺžky $n-k$, $xz \in L \iff yz \in L$.

V [5] autori dokázali nasledovné tvrdenie:

Lema 2.1. *Ak $L \in \mathsf{SPACE}(1)/f(n)$, tak pre každé n a každé $k \leq n$ relácia $\equiv_{L,n,k}$ má $O(f(n))$ tried ekvivalencie.*

Jazyk $L_{\mathsf{PAL}} = \{ww^R \mid w \in \{a,b\}^*\}$ nepatrí do $\mathsf{SPACE}(1)/f(n)$, pre žiadnu subexponenciálnu funkciu $f(n)$ [5]. Nech $L_{\mathsf{PAL2}} = \{w\#w^R \mid w \in \{a,b\}^*\}$. Zjavne $L_{\mathsf{PAL}} \in \mathsf{CFL}/0$ a $L_{\mathsf{PAL2}} \in \mathsf{DCFL}/0$.

Veta 2.3. *Jazyk $L_{\mathsf{PAL2}} = \{w\#w^R \mid w \in \{a,b\}^*\}$ nepatrí do $\mathsf{SPACE}(1)/f(n)$, pre žiadnu subexponenciálnu funkciu $f(n)$.*

Dôkaz. Pozrime sa na množinu $S = \{a,b,\#\}^k$, kde $k = \frac{n-1}{2}$, pre nepárne n, a na jej podmnožinu $S' = \{a,b\}^k$. Žiadne dve slová $x, y \in S'$ nespĺňajú $x \equiv_{L,n,k} y$. S' má 2^k prvkov, a teda $\equiv_{L,n,k}$ má aspoň 2^k tried ekvivalencie. Asymptoticky to je $\Omega(2^{n/2})$ tried ekvivalencie.

Ak by $L_{\mathsf{PAL2}} \in \mathsf{SPACE}(1)/f(n)$ pre nejakú supexponenciálnu funkciu $f(n)$, relácia $\equiv_{L,n,k}$ by mala $O(f(n))$ tried ekvivalencie, čo vedie k sporu. $\square$

Dôsledok 2.3.1. *Pre každé $k \in \mathbb{N}$, trieda $\mathsf{SPACE}(1)/k+1$ je neporovnateľná s DCFL/k, a $\mathsf{SPACE}(1)/k+1$ je tiež neporovnateľná s CFL/k.*

Veta 2.4. *Ak $f(n)$ je v $o(g(n))$ a $g(n)$ je subexponenciálna funkcia, tak existuje jazyk L taký, že L patrí do $\mathsf{DCFL}/g(n)$, ale nepatrí do $\mathsf{CFL}/f(n)$.*

Dôkaz. Dôkaz bude konštruktívny, popíšeme jazyk L a ukážeme, že spĺňa tvrdenie vety. Jazyk L bude podmnožinou $\{w\#^{n-\lfloor \log(g(n)) \rfloor} \mid n \in N,\ w \in \{a,b\}^{\lfloor \log(g(n)) \rfloor}\}$

Máme kódovanie nedeterministických zásobníkových automatov s binárnou poradnou páskou na prirodzené čísla (poradnú funkciu nekódujúc). Nezáleží na tom aké, predpokladáme nejaké fixné kódovanie. Automat, ktorý sa kóduje na číslo m budeme značiť A_m.

Pre každé $m, l \in \mathbb{N}$ si definujme jazyk $L_{m,l}$. Nech $n = 2^m(2l+1)$ (toto zaručuje rôzne n pre každú rôznu kombináciu m a l). Nech $L_{m,l}$ je taká podmnožina množiny $\{w\#^{n-\lfloor \log(g(n)) \rfloor} \mid w \in \{a,b\}^{\lfloor \log(g(n)) \rfloor}\}$ (na výber $2^{2^{\lfloor \log(g(n)) \rfloor}}$ možností), ktorá sa nezhoduje s $L(A_m, h) \cap \{a,b,\#\}^n$ pre žiadne h, pre ktoré $|h(n)| \leq 2^{\lfloor \log(g(n)) \rfloor} - 1$. Za predpokladu, že $|h(n)| \leq 2^{\lfloor \log(g(n)) \rfloor} - 1$ existuje iba $2^{2^{\lfloor \log(g(n)) \rfloor}} - 1$ možností, ako môže vyzerať množina $L(A_m, h) \cap \{a,b,\#\}^n$ (napriek tomu že možných funkcií h je nekonečne veľa, možných hodnôt $h(n)$ je konečne veľa).

Jazyk L zostrojíme zjednotením týchto jazykov:

$$L = \bigcup_{m,l=0}^{\infty} L_{m,l}$$

Teraz ukážeme, že L sa líši od každého jazyka z $\mathsf{CFL}/f(n)$. Nech $L_2 \in \mathsf{CFL}/f(n)$. Musí existovať automat A, a poradná funkcia $h : \mathbb{N} \to \{0,1\}^*$, taká že $L_2 = L(A,h)$. Nech m je kód automatu A. Keďže $|h(n)| \in o(g(n))$, musí existovať také n_0, že pre všetky $n > n_0$, $g(n) > 4\,|h(n)|$. Z konštrukcie je zjavné, že $L_2 \cap \{a,b,\#\}^{2^m(2n_0+1)} \neq L_{m,n_0}$. Avšak $L \cap \{a,b,\#\}^{2^m(2n_0+1)} = L_{m,n_0}$. Teda L a L_2 sú rôzne jazyky.

Automat A sme brali nad binárnou poradnou abecedou, každý jazyk z $\mathsf{CFL}/f(n)$ sa dá akceptovať automatom s binárnou poradnou abecedou. To preto, že ku každému neuniformnému zásobníkovému automatu, existuje ekvivalentný s binárnou abecedou (znaky z pôvodnej abecedy binárne zakódujeme, rada sa tým pádom predĺži konštantným faktorom).

Zásobníkový automat akceptujúci L bude vyzerať takto. Na poradnej páske dostane binárny strom, v ktorom listy budú znak 0 alebo 1 a podstromy budú ohraničené zátvorkami (príklad "((01)(11))((00)(10))"). Automat postupne číta písmená zo vstupu. Ak prečíta znak a vnorí sa do ľavého podstromu, ak prečíta b vnorí sa do pravého (preskočí ľavý, počíta si zátvorky na zásobníku). Keď sa dostane do listu, skontroluje, či zvyšok vstupného slova tvoria znaky #, a akceptuje ak list je 1, neakceptuje ak list je 0. Tým pádom každý list nám o inom slove hovorí, či patrí alebo nepatrí do jazyka.

Tento strom má $2^{\lfloor \log(g(m)) \rfloor}$ listov. Vnútorných vrcholov stromu je menej než listov. Pre každý vnútorný vrchol, rada obsahuje dva znaky zátvoriek. Pre každý list rada obsahuje jeden znak 0 alebo 1. Teda celá rada je dlhá najviac $3g(n)$ a teda je v $O(g(n))$.

□

Dôsledok 2.4.1. *Ak* $f(n) \in o(g(n))$, *tak* $\mathsf{CFL}/f(n) \subsetneq \mathsf{CFL}/g(n)$

Dôsledok 2.4.2. *Ak* $f(n) \in o(g(n))$, *tak* $\mathsf{DCFL}/f(n) \subsetneq \mathsf{DCFL}/g(n)$

Veta 2.5. *Jazyk* $L_{\mathsf{PAD}} = \{ww^R\#c^n \mid w \in \{a,b\}^*, n \in \mathbb{N}\}$ *nepatrí do triedy* $\mathsf{DCFL}/f(n)$, *ak* $f(n)$ *je subexponenciálna funkcia.*

Dôkaz. Povieme že znak na zásobníku korešponduje so znakom vstupného slova ak tento znak bol na zásobník pridaný keď sa čítacia hlava nachádzala na danom znaku slova. K súvislému úseku vstupného slova korešponduje súvislý úsek na zásobníku (alebo nič). V každej konfigurácii, ak sa úsek A na vstupnom slove nachádza skôr ako úsek B, tak úsek zásobníka korešpondujúci k A musí byť na zásobníku nižšie ako úsek korešpondujúci k B.

Tvrdenie dokážeme sporom: Nech A je deterministický zásobníkový automat v normálnom tvare (veta 1.1), a h je poradná funkcia pričom $|h(n)| \in O(f(n))$. Predpokladáme, že $L(A,h) = L_{\mathsf{PAD}}$.

Pri deterministických automatoch, ak existujú dva rôzne akceptačné výpočty toho istého slova, tak jeden musí byť prefixom druhého. Ak dve rôzne slová rovnakej dĺžky majú spoločný prefix, tak sa aj výpočty na týchto slovách budú zhodovať, až dokým sa hlava na vstupnej páske nedostane za ich spoločný prefix.

V dôkaze sa pozrieme na slová tvaru $w = xyy^R x^R \# c^{4n}$, kde $x, y \in \{a, b\}^n$, ktoré budeme chápať ako rozdelené na štyri úseky $w = u_1 u_2 u_3 u_4$, kde $u_1 = x$, $u_2 = y$, $u_3 = y^R x^R$, $u_4 = \# c^{4n}$, a na slová tvaru $w = xyy^R x^R xyy^R x^R \#$, kde $x, y \in \{a, b\}^n$, ktoré budeme chápať ako rozdelené na štyri úseky $w = u_1 u_2 u_3 u_4$, kde $u_1 = x$, $u_2 = y$, $u_3 = y^R x^R$, $u_4 = xyy^R x^R \#$. Každé slovo prvého tvaru sa zhoduje s jedným slovom druhého tvaru na prefixe $xyy^R x^R$ a teda aj výpočty nad týmito slovami sa musia zhodovať dokým sa čítacia hlava nedostane za tento prefix.

Rozoberieme dva prípady:

Ak pre dané n, pre aspoň polovicu možných dvojíc $x, y \in \{a, b\}^n$ sa vo výpočte nad slovom $xyy^R x^R \# c^{4n}$, keď sa čítacia hlava dostane do úseku u_4, nachádza na zásobníku časť korešpondujúca k úseku u_2, tak z Dirichletovho princípu musí existovať také $y \in \{a, b\}^n$, že pre aspoň 2^{n-1} rôznych $x \in \{a, b\}^n$ platí, že pri výpočte nad slovom $xyy^R x^R \# c^{4n}$, keď sa čítacia hlava dostane do úseku u_4, tak sa na zásobníku bude nachádzať časť korešpondujúca k úseku u_2. Označme X množinu týchto x (teda $|X| \geq 2^{n-1}$).

Pozrieme sa na dve rôzne slová $x_1 yy^R x_1^R \# c^{4n}$ a $x_2 yy^R x_2^R \# c^{4n}$, kde $x_1, x_2 \in X$. V akceptačnom výpočte automatu A nad slovom $x_1 yy^R x_1^R \# c^{4n}$ nastane konfigurácia m_{11} v ktorej sa posledný krát vyberie zo zásobníka znak korešpondujúci k úseku u_1, počas prechodu úsekom u_2. Neskôr vo výpočte nastane konfigurácia m_{12} v ktorej sa vyberie zo zásobníka posledný znak korešpondujúci k úseku u_2 počas prechodu úsekom u_4. V akceptačnom výpočte nad slovom $x_2 yy^R x_2^R \# c^{4n}$ si rovnako označíme konfigurácie m_{21}, m_{22}.

Ak sa konfigurácia m_{11} zhoduje s m_{21} v stave, pozícii na vstupnej a poradnej páske a znaku na vrchu zásobníka, a m_{12} sa tiež zhoduje s m_{22} v stave, pozícii na vstupnej a poradnej páske a znaku na vrchu zásobníka, tak A akceptuje slovo $x_1 yy^R x_2^R \# c^{4n}$ (čo by bol spor s predpokladom). Výpočet nad slovom $x_1 yy^R x_2^R \# c^{4n}$ sa bude zhodovať s výpočtom nad slovom $x_1 yy^R x_1^R \# c^{4n}$ až dokým sa nedostane do konfigurácie m_{11}, od toho momentu sa bude zhodovať s výpočtom nad slovom $x_2 yy^R x_2^R \# c^{4n}$ (až na spodnú časť zásobníka ktorej sa však nedotkne), výpočet sa potom dostane do konfigurácie m_{12} a od nej sa bude opäť zhodovať s výpočtom nad slovom $x_1 yy^R x_1^R \# c^{4n}$.

Aby nič takéto nenastalo, muselo by každé $x \in X$ mať v konfiguráciách m_{11} a m_{12} unikátnu kombináciu pozícii na vstupnej a poradnej páske, stavu a vrchného znaku na zásobníku. Týchto kombinácii je však len $O(n^2 f(8n)^2)$ čo je pre dostatočne veľké n menej ako $|X| \geq 2^{n-1}$. Teda máme spor.

Ak pre dané n, pre aspoň polovicu možných dvojíc $x, y \in \{a, b\}^n$ sa vo výpočte nad slovom $xyy^R x^R \# c^{4n}$, keď sa čítacia hlava dostane do úseku u_4, nenachádza na zásobníku časť korešpondujúca k úseku u_2, tak z Dirichletovho princípu musí existovať také $x \in \{a,b\}^n$, že pre aspoň 2^{n-1} rôznych $y \in \{a,b\}^n$ platí, že pri výpočte nad slovom $xyy^Rx^Rxyy^Rx^R\#$, keď sa čítacia hlava dostane do úseku u_4, tak sa na zásobníku už nebude nachádzať časť korešpondujúca k úseku u_2. Označme Y množinu týchto y (teda $|Y| \geq 2^{n-1}$).

Pozrieme sa na dve rôzne slová $xy_1y_1^Rx^Rxy_1y_1^Rx^R\#$ a $xy_2y_2^Rx^Rxy_2y_2^Rx^R\#$, kde $y_1, y_2 \in Y$. V akceptačnom výpočte automatu A nad slovom $xy_1y_1^Rx^Rxy_1y_1^Rx^R\#$ nastane konfigurácia m_{12} v ktorej sa prvý krát počas prechodu úsekom u_3 nachádzal na vrchu zásobníka znak korešpondujúci k úseku u_1. Počas prechodu úsekom u_1 musela nastať konfigurácia m_{11} v ktorej bol tento znak na zásobník pridaný. V akceptačnom výpočte nad slovom $xy_2y_2^Rx^Rxy_2y_2^Rx^R\#$ si rovnako označíme konfigurácie m_{21}, m_{22}.

Ak sa konfigurácia m_{11} zhoduje s m_{21} v stave, pozícii na vstupnej a poradnej páske a znaku na vrchu zásobníka, a m_{12} sa tiež zhoduje s m_{22} v stave, pozícii na vstupnej a poradnej páske a znaku na vrchu zásobníka, tak toto opäť bude viesť k tomu, že nasekaním slov $xy_1y_1^Rx^Rxy_1y_1^Rx^R\#$ a $xy_2y_2^Rx^Rxy_2y_2^Rx^R\#$ vieme vytvoriť slovo ktoré A akceptuje, ale nepatrí do jazyka L_{PAD}.

Aby nič takéto nenastalo, muselo by každé $y \in Y$ mať v konfiguráciách m_{11} a m_{12} unikátnu kombináciu pozícii na vstupnej a poradnej páske, stavu a vrchného znaku na zásobníku. Týchto kombinácii je však len $O(n^2 f(8n)^2)$ čo je pre dostatočne veľké n menej ako $|Y| \geq 2^{n-1}$. Teda máme spor. □

Dôsledok 2.5.1. *Ak $f(n) \in o(g(n))$ a $g(n)$ je subexponenciálna funkcia, tak triedy $\mathsf{DCFL}/g(n)$ a $\mathsf{CFL}/f(n)$ sú neporovnateľné.*

Pre ľubovoľnú subexponenciálnu funkciu $f(n)$ nám jazyky L_{PAL2} a L_{PAD} ukazujú separáciu medzi triedami $\mathsf{SPACE}(1)/f(n) \subsetneq \mathsf{DCFL}/f(n) \subsetneq \mathsf{CFL}/f(n)$. Toto platí aj pre rady konštantnej dĺžky: $\mathsf{SPACE}(1)/k \subsetneq \mathsf{DCFL}/k) \subsetneq \mathsf{CFL}/k$.

Veta 2.6. *Pre každé $k \in \mathbb{N}$ existuje jazyk L_k taký, že L_k patrí do $\mathsf{SPACE}(1)/n^{k+1}$, ale nepatrí do CFL/n^k.*

Dôkaz. Dôkaz bude konštruktívny, popíšeme jazyk L_k a ukážeme, že spĺňa tvrdenie vety. Jazyk L_k bude podmnožinou tejto množiny:

$$\{b^{c_1}a^{n-c_1}\#b^{c_2}a^{n-c_2}\#\ldots\#b^{c_{k+1}}a^{n-c_{k+1}}|n \in \mathbb{N}, c_i \in \{0,1,\ldots,n\}\}$$

Máme kódovanie nedeterministických zásobníkových automatov na prirodzené čísla (poradnú funkciu nekódujúc). Automat, ktorý sa kóduje na číslo m budeme značiť A_m.

Pre každé $m, l \in \mathbb{N}$ si definujme jazyk $L_{k,m,l}$. Nech $n = 2^m(2l+1)$. Nech $L_{k,m,l}$ je taká podmnožina množiny $\{b^{c_1}a^{n-c_1}\#b^{c_2}a^{n-c_2}\#\ldots\#b^{c_{k+1}}a^{n-c_{k+1}}|c_i \in \{0,1,\ldots,n\}\}$ (na výber $2^{n^{k+1}}$ možností), ktorá sa líši od $L(A_m,h) \cap \{a,b,\#\}^{kn+n+k}$ pre každé h, pre ktoré $|h(kn+n+k)| \leq n^{k+1}-1$. Za predpokladu že $|h(kn+n+k)| \leq n^{k+1}-1$, existuje len $2^{n^{k+1}}-1$ možností, ako môže vyzerať množina $L(A_m,h) \cap \{a,b,\#\}^{kn+n+k}$.

Jazyk L_k definujeme takto:

$$L_k = \bigcup_{m,l=0}^{\infty} L_{k,m,l}$$

Teraz ukážeme, že L_k sa líši od každého jazyka z CFL/n^k. Nech $L \in \mathsf{CFL}/n^k$. Musí existovať automat A, a poradná funkcia $h : \mathbb{N} \rightarrow \{0,1\}^*$, taká že $L = L(A,h)$. Nech m je kód automatu A. Keďže $|h(n)| \in O(n^k)$, musí existovať n_0 také, že pre všetky $n > n_0$, platí $|h(kn+n+k)| < n^{k+1}$. Z konštrukcie je zjavné, že $L \cap \{a,b,\#\}^{k2^m(2n_0+1)+2^m(2n_0+1)+k} \neq L_{k,m,n_0}$. Avšak $L_k \cap \{a,b,\#\}^{k2^m(2n_0+1)+2^m(2n_0+1)+k} = L_{k,m,n_0}$. Teda L a L_k sú rôzne jazyky.

Automat A sme brali nad binárnou poradnou abecedou, každý jazyk z CFL/n^k sa dá akceptovať automatom s binárnou poradnou abecedou.

Konečný automat akceptujúci L bude vyzerať takto. Na poradnej páske dostane strom hĺbky $k+1$ s koeficientom vetvenia n, v ktorom listy budú znak 0 alebo 1 a podstromy budú ohraničené zátvorkami. Automat postupne číta písmená zo vstupu. Ak prečíta znak b posunie sa na ďalšie dieťa (počíta si zátvorky v stave, strom má konštantnú hĺbku). Keď nadrazí na znak a, číta znaky a až kým nenarazí na # alebo koniec slova (overuje či je slovo v správnom formáte). Keď prečíta znak #, vnorí sa do podstromu. Ak bolo slovo v správnom formáte, tak po prečítaní celého slova zostane hlava na poradnej páske na liste stromu. Akceptuje ak list je 1 neakceptuje ak list je 0. Tým pádom každý list nám o jednom slove hovorí, či patrí alebo nepatrí do jazyka.

Dĺžka rady je priamoúmerná počtu listov. Listov je n^{k+1}. Dĺžka slova je v tomto prípade $kn+k+n$. Teda rada je dlhá $O(n^{k+1})$ v závislosti od dĺžky slova. $\square$

Dôsledok 2.6.1. *Pre každé $k \in \mathbb{N}$, trieda $\mathsf{SPACE}(1)/n^{k+1}$ je neporovnateľná s DCFL/n^k, a $\mathsf{SPACE}(1)/n^{k+1}$ je tiež neporovnateľná s CFL/n^k.*

Konečné automaty nedokážu zúžitkovať väčšiu ako polynomiálnu radu [3]. Ak jazyk L nepatrí do $\mathsf{SPACE}(1)/n^k$ pre žiadne k, tak L nie je možné akceptovať konečným automatom s radou ľubovolnej dĺžky.

Kapitola 3

Uzáverové vlastnosti

V tejto kapitole preskúmame uzáverové vlastnosti nášho modelu. Ukážeme ako sa líšia pre deterministické a nedeterministické zásobníkové automaty s poradnou páskou.

Nakoľko trieda $\mathsf{DCFL}/exp = \mathsf{CFL}/exp = \mathsf{ALL}$, tieto triedy sú uzavreté na akúkoľvek operáciu, a nebudeme sa nimi ďalej zapodievať.

Lema 3.1. *Jazyk* $\{x\#y\#x^R\#y^R \mid x, y \in \{a, b\}^*\}$ *nie je v* $\mathsf{CFL}/f(n)$*, ak* $f(n)$ *je subexponenciálna funkcia.*

Dôkaz. Tvrdenie dokážeme sporom: Nech A je nedeterministický zásobníkový automat v normálnom tvare (veta 1.1) a h je poradná funkcia, pričom $|h(n)| \in O(f(n))$, kde $f(n)$ je subexponenciálna funkcia. Predpokladáme, že automat A s funkciou h akceptuje jazyk $L = \{x\#y\#x^R\#y^R \mid x, y \in \{a, b\}^*\}$.

Uvažujme slovo $x\#y\#x^R\#y^R$ ako pozostávajúce zo štyroch úsekov rozdelené podľa znakov #. V dôkaze sa budeme pozerať len na slová tvaru $\{x\#y\#z\#w \mid x, y, z, w \in \{a, b\}^n, n \in \mathbb{N}\}$ (také v ktorých je každý úsek rovnako dlhý).

Uvažujme dve slová $x_1\#y\#x_1^R\#y^R$ a $x_2\#y\#x_2^R\#y^R$, kde $x_1, x_2, y \in \{a, b\}^n$. V akceptačnom výpočte automatu A nad slovom $x_1\#y\#x_1^R\#y^R$ nastane konfigurácia, v ktorej sa posledný krát vybral zo zásobníka znak korešpondujúci k prvej časti slova, počas prechodu druhou časťou slova. Označme pozíciu na ktorej sa v tejto konfigurácií nachádza čítacia hlava i_{11}. Neskôr môže nastať konfigurácia, v ktorej sa zo zásobníka vyberie posledný znak korešpondujúci k druhému úseku slova, počas prechodu štvrtou časťou slova. Označme pozíciu na ktorej sa v tejto konfigurácií nachádza čítacia hlava i_{12}. V akceptačnom výpočte slova $x_2\#y\#x_2^R\#y^R$ si rovnako označíme tieto pozície i_{21} a i_{22}.

Ak sa $i_{11} = i_{21}$, $i_{12} = i_{22}$, a výpočet slova $x_1\#y\#x_1^R\#y^R$ sa s výpočtom slova $x_2\#y\#x_2^R\#y^R$ zhodujú tiež v stave, vrchnom znaku na zásobníku a pozícií na poradnej páske v konfigurácií, keď bola hlava na pozícii i_{11} na vstupnej páske, a keď bola hlava na pozícií i_{12}, tak potom môžeme skonštruovať akceptačný výpočet slova $x_1\#y\#x_2^R\#y^R$. Začiatok výpočtu bude zhodný s výpočtom slova $x_1\#y\#x_1^R\#y^R$. Od pozície i_{11} po

i_{12} bude zhodný s výpočtom slova $x_2\#y\#x_2^R\#y^R$ až na spodnú časť zásobníka (ktorej sa však nedotkne). Od i_{12} po koniec slova bude výpočet opäť zhodný s výpočtom slova $x_1\#y\#x_1^R\#y^R$.

Podľa predpokladu A neakceptuje $x_1\#y\#x_2^R\#y^R$, pre žiadne rôzne x_1, x_2. Teda každé x_1 musí mať buď unikátnu voľbu týchto parametrov (i_{11}, i_{12}, stavy, znak na vrchu zásobníka, pozície na poradnej rade), alebo vo výpočte $x_1\#y\#x_1^R\#y^R$, keď sa dostane čítacia hlava do štvrtej časti slova, na zásobníku už nie je časť korešpondujúca ku druhej časti slova. Počet rôznych kombinácií týchto parametrov je $O(n^2 f(4n)^2)$, zatiaľ čo počet rôznych x je 2^n. Teda pre dostatočne veľké n, pre dané y, pre viac ako polovicu možných x, neexistuje výpočet slova $x\#y\#x^R\#y^R$, v ktorom by časť zásobníka korešpondujúca k druhej časti slova, bola prítomná keď čítacia hlava dorazí do štvrtej časti slova.

Uvažujme dve slová $x\#y_1\#x^R\#y_1^R$ a $x\#y_2\#x^R\#y_2^R$, kde $x, y_1, y_2 \in \{a, b\}^n$. V akceptačnom výpočte automatu A nad slovom $x\#y_1\#x^R\#y_1^R$ môže nastať konfigurácia, v ktorej sa na vrchu zásobníka prvý krát nachádzal znak korešpondujúci k prvej časti slova, počas prechodu treťou časťou slova. Označme pozíciu na ktorej sa v tejto konfigurácií nachádza čítacia hlava i_{12}. Ak toto nastalo, tak počas prechodu prvou časťou slova musela nastať konfigurácia, v ktorej bol tento znak na vrch zásobníka pridaný. Označme pozíciu na ktorej sa v tejto konfigurácií nachádza čítacia hlava i_{11}. V akceptačnom výpočte slova $x\#y_2\#x^R\#y_2^R$ si rovnako označíme tieto pozície i_{21} a i_{22}.

Ak sa $i_{11} = i_{21}$, $i_{12} = i_{22}$, a výpočet slova $x\#y_1\#x^R\#y_1^R$ sa s výpočtom slova $x\#y_2\#x^R\#y_2^R$ zhodujú tiež v stave, vrchnom znaku na zásobníku a pozícií na poradnej páske v konfigurácií, keď bola hlava na pozícii i_{11} na vstupnej páske, a keď bola hlava na pozícií i_{12}, tak potom môžeme skonštruovať akceptačný výpočet slova $x\#y_2\#x^R\#y_1^R$. Začiatok výpočtu bude zhodný s výpočtom slova $x\#y_1\#x^R\#y_1^R$. Od pozície i_{11} po i_{12} bude zhodný s výpočtom slova $x\#y_2\#x^R\#y_2^R$ až na spodnú časť zásobníka (ktorej sa však nedotkne). Od i_{12} po koniec slova bude výpočet opäť zhodný s výpočtom slova $x\#y_1\#x^R\#y_1^R$.

Podľa predpokladu A neakceptuje $x\#y_2\#x^R\#y_1^R$, pre žiadne rôzne y_1, y_2. Toto nám obmedzuje počet rôznych y_1 pre ktoré existuje výpočet slova $x\#y_1\#x^R\#y_1^R$ v ktorom by sa počas prechodu treťou časťou slova na vrchu zásobníka nachádzal znak korešpondujúci k prvej časti slova na $O(n^2 f(4n)^2)$ možných. Teda pre dostatočne veľké n, pre dané x, pre viac ako polovicu možných y, neexistuje výpočet slova $x\#y\#x^R\#y^R$, v ktorom by časť zásobníka korešpondujúca k druhej časti slova, nebola prítomná keď čítacia hlava dorazí do štvrtej časti slova.

Z Dirichletovho princípu teda musí existovať taká dvojica $x, y \in \{a, b\}^n$, že pre slovo $x\#y\#x^R\#y^R$ neexistuje ani výpočet v ktorom by časť zásobníka korešpondujúca k druhej časti slova, bola prítomná keď čítacia hlava dorazí do štvrtej časti slova, ani výpočet v ktorom by bola prítomná. Teda neexistuje žiaden akceptačný výpočet slova $x\#y\#x^R\#y^R$, čo je spor. $\square$

Veta 3.1. *Ak $f(n)$ je subexponenciálna funkcia, triedy* $\mathsf{CFL}/f(n)$ *a* $\mathsf{DCFL}/f(n)$ *nie sú uzavreté na prienik jazykov.*

Dôkaz. Uvažujme jazyky $L_1 = \{x\#y\#x^R\#z \mid x, y, z \in \{a, b\}^*\}$ a $L_2 = \{x\#y\#z\#y^R \mid x, y, z \in \{a, b\}^*\}$. Jazyky L_1 a L_2 sú deterministické bezkontextové jazyky, teda patria aj do $\mathsf{CFL}/f(n)$ a $\mathsf{DCFL}/f(n)$. Ich prienik je jazyk z lemy 3.1 a teda nepatrí do $\mathsf{CFL}/f(n)$. Keďže $\mathsf{DCFL}/f(n) \subseteq \mathsf{CFL}/f(n)$, Ich prienik nepatrí ani do $\mathsf{DCFL}/f(n)$. □

Veta 3.2. *Pre každú funkciu $f(n)$, trieda* $\mathsf{DCFL}/f(n)$ *je uzavretá na komplement.*

Ak deterministický zásobníkový automat A s poradnou funkciou h, kde $|h(n)| \in O(f(n))$, akceptujú jazyk L, je možné skonštruovať deterministický zásobníkový automat A' taký, že $L(A', h) = L^C$. Konštrukcia by bola v princípe rovnaká ako v prípade klasických (bez pomocnej pásky) deterministických zásobníkových automatov (detekcia zacyklenia pomocou počítadla v stave), nebudeme ju tu uvádzať.

Veta 3.3. *Pre každú funkciu $f(n)$, trieda* $\mathsf{CFL}/f(n)$ *je uzavretá na zjednotenie jazykov.*

Dôkaz. Nech L_1 je jazyk akceptovaný automatom A_1 s poradnou funkciou $h_1 : \mathbb{N} \to \Delta_1^*$, kde $|h_1(n)| \in O(f(n))$ a L_2 je jazyk akceptovaný automatom A_2 s poradnou funkciou $h_2 : \mathbb{N} \to \Delta_2^*$, kde $|h_2(n)| \in O(f(n))$.

Definujeme si poradnú funkciu $h : \mathbb{N} \to (\Delta_1 \cup \Delta_2 \cup \{\#\})^*$ ako $h(n) = h_1(n)\#h_2(n)$ (novozavedený znak $\#$ nepatrí do abecied Δ_1 a Δ_2). Automat A sa najprv nedeterministicky rozhodne, či bude simulovať automat A_1, alebo A_2. Ak sa rozhodne pre A_1, bude pri simulácii považovať symbol $\#$ na poradnej páske za znak značiaci koniec rady. Ak sa rozhodne odsimulovať A_2, najprv sa na poradnej páske presunie za znak $\#$ (časť za týmto znakom je rada pre A_2).

Práve vtedy ak $w \in L_1$ existuje výpočet v ktorom A akceptuje w, pričom sa na začiatku rozhodne simulovať automat A_1. Rovnako existuje výpočet v ktorom A akceptuje w, pričom sa na začiatku rozhodne simulovať automat A_2 práve vtedy keď $w \in L_2$. Preto automat A spolu s poradnou funkciou h akceptuje jazyk $L_1 \cup L_2$. □

Prienik dvoch jazykov vieme vyjadriť pomocou komplementu a zjednotenia: $L_1 \cap L_2 = (L_1^C \cup L_2^C)^C$. Ak by nejaká trieda jazykov bola uzavretá aj na zjednotenie, aj na komplement, musela by byť nutne uzavretá aj na prienik. Ako dôsledok predchádzajúcich tvrdení teda musia platiť aj nasledovné dve vety.

Veta 3.4. *Ak $f(n)$ je subexponenciálna funkcia, trieda* $\mathsf{CFL}/f(n)$ *nie je uzavretá na komplement.*

Veta 3.5. *Ak $f(n)$ je subexponenciálna funkcia, trieda* $\mathsf{DCFL}/f(n)$ *nie je uzavretá na zjednotenie jazykov.*

Veta 3.6. *Triedy* $\mathsf{DCFL}/f(n)$ *a* $\mathsf{CFL}/f(n)$ *sú uzavreté na prienik s regulárnym jazykom.*

Dôkaz. Nech A_1 je nedeterministický zásobníkový automat a h je poradná funkcia, pričom $|h(n)| \in O(f(n))$. Nech automat A_1 spolu s poradnou funkciou h akceptujú jazyk $L_1 = L(A_1, h)$. Nech A_2 je deterministický konečný automat akceptujúci jazyk L_2.

Je možné skonštruovať automat A taký, že $L(A, h) = L_1 \cap L_2$. Stavy automatu A budú mať dve zložky $K_A = K_{A_1} \times K_{A_2}$. Automat A bude súčasne simulovať aj A_1 aj A_2. Akceptačné stavy budú $F_A = F_{A_1} \times F_{A_2}$, teda akceptujeme len ak sú oba simulované automaty v akceptačnom stave.

V prvej zložke si A pamätá stav automatu A_1. Simulovaný automat A_1 rozhoduje o pohybe na oboch páskach a dianí na zásobníku.

Druhá zložku tvorí stav automatu A_2. Mení sa len keď sa má čítacia hlava posunúť dopredu. Tým pádom keď sa bude čítacia hlava nachádzať na koncovom znaku \$, bude v tejto zložke taký stav, v akom by bol A_2 po prečítaní celého slova.

Ak A_1 bol deterministický, tak vďaka takejto konštrukcii je aj A deterministický. □

Veta 3.7. *Ak dva jazyky L_1, L_2 patria do $\mathsf{CFL}/f(n)$ tak ich zreťazenie L_1L_2 patrí do $\mathsf{CFL}/n^2f(n)$.*

Dôkaz. Ukážeme konštrukciu automatu A a poradnej funkcie h takých, že $L(A, h) = L_1L_2$.

Nech L_1 je jazyk akceptovaný automatom A_1 s poradnou funkciou $h_1 : \mathbb{N} \to \Delta_1^*$, kde $|h_1(n)| \in O(f(n))$ a L_2 je jazyk akceptovaný automatom A_2 s poradnou funkciou $h_2 : \mathbb{N} \to \Delta_2^*$, kde $|h_2(n)| \in O(f(n))$. Bez ujmi na všeobecnosti, môžeme predpokladať, že poradné funkcie h_1 a h_2 vždy produkujú rady párnej dĺžky.

Definujeme si funkciu $h_1' : \mathbb{N} \to (\Delta_1 \cup \{\#_m, \#_e\})^*$. Nech w_1 a w_2 je prvá a druhá polovica slova $h_1(n)$, teda $h_1(n) = w_1w_2$ a $|w_1| = |w_2|$.

$$h_1'(n) = \underbrace{w_1\#_m w_2\#_e w_1\#_m w_2\#_e \dots \#_e w_1\#_m w_2}_{n\text{-krát}}$$

Obdobným spôsobom definujeme funkciu h_2'. Teraz môžme definovať našu poradnú funkciu $h : \mathbb{N} \to (\Delta_1 \cup \Delta_2 \cup \{\#_m, \#_e, \#_s, \#_b\})^*$. Predpokladáme že novo zavedené znaky $\#_m, \#_e, \#_s, \#_b$ sa nenachádzajú v poradných abecedách Δ_1 a Δ_2.

$$h(n) = h_1'(1)\#_s h_1'(2)\#_s \dots \#_s h_1'(n)\#_b h_2'(1)\#_s h_2'(2)\#_s \dots \#_s h_2'(n)$$

Každé slovo w z jazyka L sa skladá z dvoch častí $w = w_1w_2$, kde $w_1 \in L_1$ a $w_2 \in L_2$. Automat A sa nedeterministicky rozhodne, kde sa nachádza predel medzi týmito slovami. Následne odsimuluje výpočet automatu A_1 nad časťou slova pred týmto predelom, a potom odsimuluje výpočet automatu A_2 na zvyšku slova. Oba musia akceptovať aby A akceptoval.

Najprv sa hlava na poradnej páske presunie na začiatok niektorého z úsekov $h'_1(i)$ (niekoľkokrát sa nedeterministicky rozhodne, či sa presunie na nasledujúci znak $\#_s$ na poradnej páske, ak narazí na $\#_b$, neakceptuje). Následne začne simulovať automat A_1. Pamätá si stav automatu A_1 v stave. Ak sa A_1 chce posunúť dopredu na poradnej páske, A sa posunie dopredu na poradnej páske (preskočí $\#_m$, ak naň narazí). Ak sa A_1 chce posunúť dopredu na vstupnej páske, A sa posunie dopredu na vstupnej páske, a posunie sa na ďalší segment v $h'_1(i)$, na to isté miesto v $h_1(i)$. K tomu mu pomôžu symboly $\#_m$ a $\#_e$. Automat sa bude posúvať na poradnej páske dokým nenarazí na znak $\#_m$ alebo $\#_e$, pričom pre každý krok pridá na zásobník pomocný symbol B. Potom sa automat posunie dopredu na poradnej páske za každý symbol B, vyberajúc všetky B zo zásobníka. Následne tento proces automat zopakuje druhý krát. Ak narazí na znak $\#_s$ alebo $\#_b$ na poradnej páske, znamená to, že A_1 narazil na koniec svojej časti slova na vstupnej páske (v simulácii A_1 prečíta koncový znak \$). Keď skončí simulácia automatu A_1, hlava na poradnej páske sa presunie za symbol $\#_b$. Následne zopakuje takú istú procedúru na druhej polovici rady. Nedeterministicky sa presunie na začiatok nejakého úseku $h'_2(i)$, a rovnako odsimuluje A_2.

Automat A akceptuje iba ak obe simulácie akceptovali, a po skončení druhej simulácie skončila hlava na vstupnej páske presne na konci (ak koniec "prestrelí", t.j. počas druhej simulácie narazí na koniec vstupu tak neakceptuje). Preto A akceptačné w iba ak $w \in L_1L_2$.

Ak $w \in L_1L_2$, tak $w = w_1w_2$, kde $w_1 \in L_1$ a $w_2 \in L_2$. Automat A so vstupom w si môže nedeterministicky zvoliť úseky rady $h'_1(|w_1|)$ a $h'_2(|w_2|)$, čím sa zaručí, že A akceptuje.

Funkcie $|h'_1(n)|$ a $|h'_2(n)|$ sú $O(nf(n))$. Teda $|h(n)|$ patrí do $O\big(n\sum_{i=1}^{n} f(i)\big) \subseteq O(n^2 f(n))$. $\square$

Dôsledok 3.7.1. *Trieda* $\mathsf{CFL}/poly$ *je uzavretá na zreťazenie.*

Veta 3.8. *Ak jazyk L patrí do $\mathsf{CFL}/f(n)$ tak L^* patrí do $\mathsf{CFL}/f(n)n^2\log n$.*

Dôkaz. Nech L je jazyk akceptovaný automatom A s poradnou funkciou $h : \mathbb{N} \to \Delta^*$, kde $|h(n)| \in O(f(n))$. Ukážeme konštrukciu automatu A' a poradnej funkcie h'', ktorý akceptuje jazyk L^*. Bez ujmy na všeobecnosti, môžeme predpokladať, že poradná funkcia h vždy produkuje rady párnej dĺžky.

Uvažujme postupnosť $(a_i)_{i\in\mathbb{N}}$, ktorá zaradom vymenuje všetkých deliteľov čísla n, pre všetky čísla n. [1]

$$(a_i)_{i\in\mathbb{N}} = (1, 1, 2, 1, 3, 1, 2, 4, 1, 5, 1, 2, 3, 6, \ldots)$$

[1] Postupnosť A027750 v On-Line Encyclopedia of Integer Sequences

Nech b_n značí index prvého výskytu čísla n v postupnosti $(a_i)_{i\in\mathbb{N}}$, t.j. najmenšie číslo, pre ktoré $a_{b_n} = n$. Postupnosť $(b_i)_{i\in\mathbb{N}}$ je $O(n \log n)$, nakoľko každé n má najviac $\log n$ deliteľov.

Funkciu h' si definujeme rovnako ako v dôkaze vety 3.7. Definujeme si poradnú funkciu $h'' : \mathbb{N} \to (\Delta \cup \{\#_m, \#_e, \#_s\})^*$:

$$h''(n) = h'(a_1)\#_s h'(a_2)\#_s \dots \#_s h'(a_{b_n})$$

Automat A' bude pracovať takto: niekoľkokrát sa presunie za najbližší znak $\#_s$ na poradnej páske (po každom presunutí sa nedeterministicky rozhodne, či pokračuje). Zastaví sa na začiatku nejakého úseku $h'(i)$. Následne odsimuluje výpočet automatu A na časti vstupu dĺžky i. Simulácia bude prebiehať rovnako ako tá v dôkaze vety 3.7. Po skončení simulácie skončí hlava na poradnej páske za úsekom $h'(i)$, a hlava na vstupnej páske skončí o i krokov ďalej. Toto budeme opakovať v cykle (nedeterministicky sa presunúť na niektorý úsek vstupnej pásky, a potom ho použiť na simuláciu A na ďalšom úseku vstupu). Nakoniec A' akceptuje iba ak všetky simulované automaty A akceptovali svoj úsek vstupu, a cyklus skončí presne na konci vstupu (ak koniec "prestrelí", t.j. počas simulácie narazí na koniec vstupu tak neakceptuje).

Z konštrukcie je evidentné, že ak existuje akceptačný výpočet kde A' akceptuje slovo w, tak w sa musí dať rozdeliť na niekoľko slov $w = w_1 w_2 \dots w_m$ tak, že $w_i \in L$ pre všetky i.

Postupnosť $(a_i)_{i=0}^{b_n}$ má takúto vlastnosť: pre ľubovoľné kladné celé čísla $c_1, c_2, \dots, c_m$ také, že $\sum_{i=1}^{m} c_i = n$, postupnosť $(c_i)_{i=1}^{m}$ je podpostupnosť postupnosti $(a_i)_{i=0}^{b_n}$. Preto ak slovo w patrí do L^*, teda $w = w_1 w_2 \dots w_m$, kde všetky $w_i \in L$, tak musí existovať výpočet ktorý w akceptuje. Garantovane si A' vie nedeterministicky zvoliť také úseky rady, ktoré správne korešpondujú k dĺžkam úsekov w_i.

Funkcia $|h'(n)|$ je $O(nf(n))$. Teda $|h''(n)|$ patrí do $O(f(n)n^2 \log n)$. □

Dôsledok 3.8.1. *Trieda* $\mathsf{CFL}/poly$ *je uzavretá na Kleeneho uzáver.*

Veta 3.9. *Ak $f(n)$ je subexponenciálna funkcia a $f(n) \in \Omega(n)$, trieda* $\mathsf{DCFL}/f(n)$ *nie je uzavretá na zreťazenie jazykov.*

Dôkaz. Jazyk $L_{\mathsf{PAL}} = \{ww^R \mid w \in \{a,b\}^*\}$ je možné akceptovať s lineárnou radou. Pre slová párnej dĺžky bude rada dlhá polovicu dĺžky slova. Hlava na poradnej páske sa bude pohybovať zarovno s hlavou na vstupnej páske. Koniec poradnej pásky značí polovicu slova. Automat znaky zo vstupu pridáva na zásobník kým nepríde do stredu slova. Pri prechode druhou polovicou slova znaky zo zásobníka vyberá a kontroluje či sa zhodujú so vstupom. Slová nepárnej dĺžky neakceptuje.

Zreťazenie s regulárnym jazykom $\{\#c^n \mid n \in \mathbb{N}\}$ nám dáva jazyk $L_{\mathsf{PAD}} = \{ww^R\#c^n \mid w \in \{a,b\}^*, n \in \mathbb{N}\}$. Tento jazyk nepatrí do $\mathsf{DCFL}/f(n)$ (veta 2.5). □

Dôsledok 3.9.1. *Trieda* $\mathsf{DCFL}/poly$ *nie je uzavretá na zreťazenie.*

Veta 3.10. *Ak* $f(n)$ *je subexponenciálna funkcia a* $f(n) \in \Omega(n)$, *trieda* $\mathsf{DCFL}/f(n)$ *nie je uzavretá na Kleeneho uzáver.*

Dôkaz. Jazyk $L = \{ww^R\# \mid w \in \{a,b\}^*\} \cup \{c^n \mid n \in \mathbb{N}\}$ je možné akceptovať s lineárnou radou. Rada bude opäť značiť stred slova. Ak je prvý znak vstupu c, automat len skontroluje či sú všetky znaky c. Inak automat postupuje rovnako, ako automat v predchádzajúcom dôkaze, navyše kontroluje, že žiaden znak nie je c.

Prienik jazyka L^* a regulárneho jazyka $\{w\#c^n \mid w \in \{a,b\}^*, n \in \mathbb{N}\}$ nám dáva $L_{\mathsf{PAD}} = \{ww^R\#c^n \mid w \in \{a,b\}^*, n \in \mathbb{N}\}$. Jazyk L_{PAD} nepatrí do $\mathsf{DCFL}/f(n)$ (veta 2.5). Nakoľko $\mathsf{DCFL}/f(n)$ je uzavretá na prienik s regulárnym jazykom, nemôže byť uzavretá na Kleeneho uzáver. □

Dôsledok 3.10.1. *Trieda* $\mathsf{DCFL}/poly$ *nie je uzavretá na Kleeneho uzáver.*

Otázka zreťazenia pre triedy $\mathsf{DCFL}/f(n)$, kde funkcia $f(n) \in o(n)$ zostáva otvorená.

Záver

V práci sme preskúmali vlastnosti neuniformného modelu zásobníkových automatov, ktoré dostávajú pomocnú informáciu závislú od dĺžky vstupného slova na dodatočnej páske.

Pre radu subexponenciálnej dĺžky sme ukázali existenciu hierarchií tried jazykov oboch variantov (deterministický a nedeterministický) nášho modelu v závislosti od dĺžky rady. Využili sme na to diagonalizačné dôkazy. Ukázali sme tiež, že pri radách rovnakej dĺžky existuje istá forma Chomského hierarchie, t.j. trieda nedeterministických zásobníkových automatov s radou dĺžky $O(f(n))$ je väčšia od triedy deterministických zásobníkových automatov s radou dĺžky $O(f(n))$ a tá je zas väčšia od triedy deterministických konečných automatov s radou dĺžky $O(f(n))$, pre subexponenciálne funkcie f.

Uviedli sme niekoľko výsledkov o uzáverových vlastnostiach tohto modelu. Tieto výsledky sa líšia medzi deterministickým a nedeterministickým variantom. Zatiaľčo trieda nedeterministických zásobníkových automatov s radou polynomiálnej dĺžky je uzavretá na zjednotenie, zreťazenie a Kleeneho uzáver, trieda deterministických zásobníkových automatov s radou polynomiálnej dĺžky je uzavretá iba na komplement (nie je uzavretá na prienik, zjednotenie, zreťazenie, ani Kleeneho uzáver). Toto sa zhoduje s tým ako tomu je aj pre klasických zásobníkových automatoch (bez pomocnej informácie).

Literatúra

[1] C. Damm and M. Holzer. Automata that take advice. *LNCS*, 969:149--158, 1995.

[2] Pavol Ďuriš, Rafael Korbaš, Rastislav Královič, and Richard Královič. Determinism and nondeterminism in finite automata with advice. In *Adventures Between Lower Bounds and Higher Altitudes: Essays Dedicated to Juraj Hromkovič on the Occasion of His 60th Birthday*, pages 3--16. Springer, 2018.

[3] Pavol Ďuriš, Rastislav Královič, Richard Královič, Dana Pardubská, Martin Pašen, and Peter Rossmanith. Randomization in non-uniform finite automata. In *45th International Symposium on Mathematical Foundations of Computer Science (MFCS 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.

[4] Richard Karp. Turing machines that take advice. *Enseign. Math.*, 28:191--209, 1982.

[5] UĞUR KÜÇÜK and A. C. CEM SAY. Finite automata with advice tapes. *International Journal of Foundations of Computer Science*, 25(8):987--1000, 2014.