

Supplementary Information and Transducers

Matúš Juran

Vedúci práce: prof. RNDr. Branislav Rován, PhD.

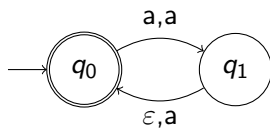
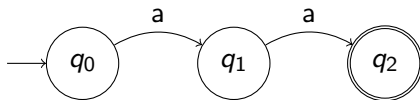
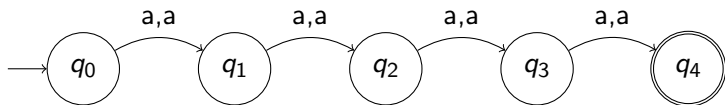
FMFI UK

26.6.2019

Motivácia

- Skúmanie užitočnosti informácie
- Formalizácia - dodatočná informácia ako predpoklad o vstupnom jazyku
- Predchádzajúce práce - prieniky jazykov
- Vida - možnosť transformácie dodatočnej informácie prekladačom
- Naša práca - zjednodušenie “výroby” jazyka

Príklad



Definície

Definícia

Deterministický sekvenčný prekladač je 6-tica

$M = (K, \Sigma_1, \Sigma_2, H, q_0, F)$ kde

$H \subseteq K \times (\Sigma_1 \cup \{\varepsilon\}) \times (\Sigma_2 \cup \{\varepsilon\}) \times K$ a ďalší krok výpočtu je vždy jednoznačne určený.

Definícia

Nedeterministický sekvenčný prekladač je 6-tica

$M = (K, \Sigma_1, \Sigma_2, H, q_0, F)$ kde

$H \subseteq K \times (\Sigma_1 \cup \{\varepsilon\}) \times (\Sigma_2 \cup \{\varepsilon\}) \times K$.

Definície

Definícia

Deterministická stavová zložitosť jazyka L je definovaná ako
 $sc(L) = \min\{|K_A| \mid A \text{ je DKA, } L(A) = L\}$.

Definícia

Nedeterministická stavová zložitosť jazyka L je definovaná ako
 $nsc(L) = \min\{|K_A| \mid A \text{ je NKA, } L(A) = L\}$.

Definície

Definícia

Deterministická prekladačová zložitosť jazyka L je definovaná ako $tsc(L) = \min\{|K_M| \mid M(\Sigma_L^*) = L, M \text{ je deterministický prekladač}\}$.

Definícia

Nedeterministická prekladačová zložitosť jazyka L je definovaná ako $ntsc(L) = \min\{|K_M| \mid M(\Sigma_L^*) = L, M \text{ je nedeterministický prekladač}\}$.

Definície

Definícia

Nech $L_{adv} \in \mathcal{R}$. Trieda $\mathcal{L}_D(L_{adv})$ je definovaná ako

$$\mathcal{L}_D(L_{adv}) = \{L \mid sc(L_{adv}) < tsc(L), \exists M : M \text{ je deterministický prekladač, } M(L_{adv}) = L, |\Sigma_{L_{adv}}| \leq |\Sigma_L|, \|K_M\| < tsc(L)\}$$

Definícia

Nech $L_{adv} \in \mathcal{R}$. Trieda $\mathcal{L}_N(L_{adv})$ je definovaná ako

$$\mathcal{L}_N(L_{adv}) = \{L \mid nsc(L_{adv}) < ntsc(L), \exists M : M \text{ je nedeterministický prekladač, } M(L_{adv}) = L, |\Sigma_{L_{adv}}| \leq |\Sigma_L|, |K_M| < ntsc(L)\}$$

Definície

Definícia

Trieda \mathcal{L}_{DTA} je definovaná ako

$$\mathcal{L}_{DTA} = \{L \mid \exists L_{adv} : L \in \mathcal{L}_D(L_{adv})\}.$$

Definícia

Trieda \mathcal{L}_{NTA} je definovaná ako

$$\mathcal{L}_{NTA} = \{L \mid \exists L_{adv} : L \in \mathcal{L}_N(L_{adv})\}.$$

Techniky

Definícia

Nech $L \in \mathcal{R}$, nech $P = \{(x_i, y_i) \mid 1 \leq i \leq n, x_i y_i \in L\}$. Ak $i \neq j$, nech $x_i y_j \notin L$ alebo $x_j y_i \notin L$. Potom P je rozšírená mätúca množina pre L .

Veta

Nech $L \in \mathcal{R}$, nech P je rozšírená mätúca množina pre L . Potom $nsc(L) \geq |P|$.

- existujúca technika

Techniky

Veta

Nech $L \in \mathcal{R}$. Potom $tsc(L) \geq ntsc(L) = nsc(L)$.

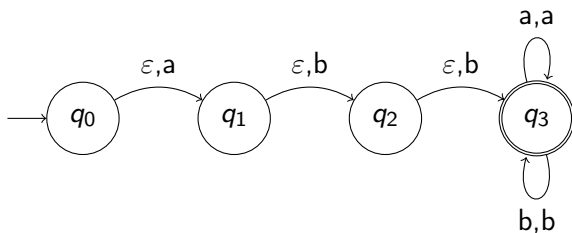
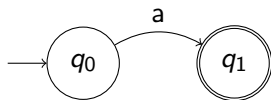
Jazyky v \mathcal{L}_{NTA}

Veta

Nech $|w| \geq 2$. Potom $\{w\} \in \mathcal{L}_{NTA}$.

Jazyky v \mathcal{L}_{NTA}

$$L = \{abba\}$$



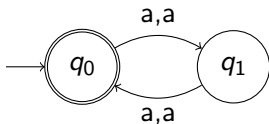
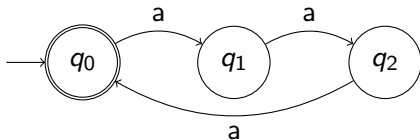
Jazyky v \mathcal{L}_{NTA}

Veta

Nech n je zložené číslo. Potom $\{a^{kn} \mid k \in \mathbb{N}\} \in \mathcal{L}_{NTA}$.

Jazyky v \mathcal{L}_{NTA}

$$L = \{a^{6n} \mid n \in \mathbb{N}\}$$



Jazyky mimo \mathcal{L}_{NTA}

Veta

Ak $ntsc(L) \leq 2$, potom $L \notin \mathcal{L}_{NTA}$.

Jazyky mimo \mathcal{L}_{NTA}

Veta

Nech p je prvočíslo. Potom $\{a^{kp} \mid k \in \mathbb{N}\} \notin \mathcal{L}_{NTA}$.

Jazyky mimo \mathcal{L}_{NTA}

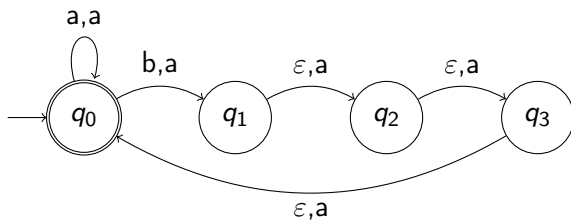
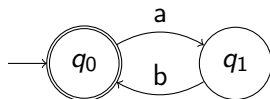
Veta

Nech p je prvočíslo. Potom $\{a^{kp} \mid k \in \mathbb{N}\} \notin \mathcal{L}_{NTA}$.

- Ak by Σ_2 obsahovala znak b , jazyk $L_{adv} = \{ab\}^*$ by bol užitočnou informáciou

Jazyky mimo \mathcal{L}_{NTA}

$$L = \{a^{5n} \mid n \in \mathbb{N}\}$$



Jazyky v \mathcal{L}_{DTA}

Veta

Nech $|w| \geq 3$. Potom $\{w\} \in \mathcal{L}_{DTA}$.

Jazyky v \mathcal{L}_{DTA}

Veta

Nech n je zložené číslo. Potom $\{a^{kn} \mid k \in \mathbb{N}\} \in \mathcal{L}_{DTA}$.

Jazyky mimo \mathcal{L}_{DTA}

Veta

Ak $tsc(L) \leq 2$, potom $L \notin \mathcal{L}_{DTA}$.

Jazyky mimo \mathcal{L}_{DTA}

Veta

Nech p je prvočíslo. Potom $\{a^{kp} \mid k \in \mathbb{N}\} \notin \mathcal{L}_{DTA}$.

Jazyky mimo \mathcal{L}_{DTA}

Veta

Nech p je prvočíslo. Potom $\{a^{kp} \mid k \in \mathbb{N}\} \notin \mathcal{L}_{DTA}$.

- Ak by Σ_2 obsahovala znak b , jazyk $L_{adv} = \{ab\}^*$ by bol užitočnou informáciou

Vzťah s rozložiteľnosťou automatov

Veta

Existuje deterministicky nerozložiteľný jazyk L , ktorý patrí do triedy \mathcal{L}_{DTA} .

Veta

Trieda nedeterministicky rozložiteľných jazykov je vlastnou podtriedou triedy \mathcal{L}_{NTA} .

Veta

Ak je jazyk L deterministicky rozložiteľný a $sc(L) = tsc(L)$, potom $L \in \mathcal{L}_{DTA}$

Vzťah s rozložiteľnosťou automatov

Hypotéza

Deterministicky rozložiteľné jazyky a trieda \mathcal{L}_{DTA} sú neporovnateľné

- $L = c(a^2\{b\}^*\{a, b\})^*$
- $sc(L) \geq 7$
- $tsc(L) \leq 4$
- Hypotéza je overiteľná hrubou silou v konečnom čase

Vzťah \mathcal{L}_{DTA} a \mathcal{L}_{NTA}

Veta

$$\mathcal{L}_{NTA} \not\subseteq \mathcal{L}_{DTA}$$

- $L = \{a^2\}$
- Dokázané rozborom prípadov

Uzáverové vlastnosti

	h_ϵ	h^{-1}	*	+	\cap	\cup	.	bij. h
\mathcal{L}_{DTA}	X	X	X	X	X	X	X	✓
\mathcal{L}_{NTA}	X	X	X	X	X	X	X	✓

- Príklad: $\{a^{4n} \mid n \in \mathbb{N}\} \cup \{a^{4n+2} \mid n \in \mathbb{N}\} = \{a^{2n} \mid n \in \mathbb{N}\}$

Fixná informácia

	h_ϵ	h^{-1}	*	+	\cap	\cup	.	C	bij. h
$\mathcal{L}_D(L_{adv})$	X	X	X	X	X	X	X	X	✓
$\mathcal{L}_N(L_{adv})$	X	X	X	X	X	X	?	?	✓

- Ak trieda nie je uzavretá na unárnu operáciu, nemusia byť na ňu uzavreté ani triedy generované fixnou informáciou

Modifikácie informácie

	h	R	\cap	\cup	bij. h
$\mathcal{L}_D(L_{adv})$	\times	\times	\times	\times	\checkmark

- Príklad:
- $\{a^{10n} \mid n \in \mathbb{N}\} \in \mathcal{L}_D(\{a^{2n} \mid n \in \mathbb{N}\})$
- $\{a^{10n} \mid n \in \mathbb{N}\} \in \mathcal{L}_D(\{a^{2n+1} \mid n \in \mathbb{N}\})$
- $\{a^{2n} \mid n \in \mathbb{N}\} \cup \{a^{2n+1} \mid n \in \mathbb{N}\} = \{a\}^*$

Zhrnutie

- Existujú zjednodušiteľné nerozložiteľné jazyky
- Nedeterministicky rozložiteľné sú vlastná podtrieda \mathcal{L}_{NTA}
- Na veľkosti abecedy záleží
- Premenovanie znakov zachováva užitočnosť
- \mathcal{L}_{DTA} aj \mathcal{L}_{NTA} majú zlé uzáverové vlastnosti
- Triedy generované fixnou informáciou môžu mať zlé uzáverové vlastnosti
- Úpravy môžu spraviť informáciu neužitočnou

Možnosti pokračovania výskumu

- Overenie hypotézy o deterministicky rozložiteľných jazykoch a \mathcal{L}_{DTA}
- Skúmanie zvyšných uzáverových vlastností
- Charakterizácia tried \mathcal{L}_{DTA} a \mathcal{L}_{NTA}
- Skúmanie množstva ušetriteľných stavov

Ďakujem za pozornosť