Języki, automaty i obliczenia

Wykład 8: Obrazy przemienne języków. Automaty dwukierunkowe.

Sławomir Lasota

Uniwersytet Warszawski

24 kwietnia 2024

Plan

Obraz przemienny języka

- Dla $w = a_1 \ldots a_n \in A^*$, zdefiniujmy funkcję (multizbiór nad A)

$$\mathcal{P}(w) : A \to \mathbb{N} \qquad \mathcal{P}(w)(a) = |\{ i \in \{1 \ldots n\} : a_i = a \}|.$$

- *Obraz przemienny* języka $L \subseteq A^*$ to:

$$\mathcal{P}(L) = \{ \mathcal{P}(w) : w \in L \}.$$

- Dla ustalonego liniowego porządku na alfabecie $A = \{b_1, \ldots b_k\}$,

$$b_1 < b_2 < \ldots < b_k,$$

  obraz przemienny języka L można utożsamiać z podzbiorem $\mathbb{N}^k$:

$$\mathcal{P}(w)(i) = \mathcal{P}(w)(b_i).$$

Przykład

Niech $A = \{a, b\}, \; a < b$.

$$\mathcal{P}(\{ a^n b^n a : n \in \mathbb{N} \}) = \mathcal{P}((a\,b)^*\,a) = \{ (n+1, n) : n \in \mathbb{N} \}$$
$$\mathcal{P}(\text{palindromy}) = \{ (n, m) : \text{przynajmniej jedno spośród } n, m \text{ parzyste} \}$$

Zbiory liniowe

Dla $P \subseteq \mathbb{N}^k$, zdefiniujmy $P^\oplus = \{\, p_1 + p_2 + \ldots + p_m \;:\; m \geq 0,\; p_1, p_2, \ldots, p_m \in P \,\}$.

Zbiór *liniowy* $X \subseteq \mathbb{N}^k$ to dowolny zbiór postaci

$$X = b + P^\oplus = \{\, b + p_1 + \ldots + p_m \;:\; p_1, \ldots, p_m \in P \,\}, \qquad \text{dla } b \in \mathbb{N}^k,\; P \subseteq_{\text{fin}} \mathbb{N}^k.$$

$$= \{\, b + p \;:\; b \in B,\; p \in P^\oplus \,\}$$

b baza, P okresy

Przykład

Niech $b = (1,2)$, $P = \{(2,0),(0,3)\}$.

$$b + P^\oplus = \{\, (n,m) \in \mathbb{N}^2 \;:\; n \equiv 1 \,(\mathrm{mod} 2),\; m \equiv 2 \,(\mathrm{mod} 3) \,\}.$$

rysunek

Przykład

Niech $b = (0,0)$, $P = \{(1,1),(0,3)\}$.

$$b + P^\oplus = \{\, (n,m) \in \mathbb{N}^2 \;:\; 0 \leq m - n \equiv 0 \,(\mathrm{mod} 3) \,\}.$$

rysunek

Zbiory semiliniowe

Dla $P \subseteq \mathbb{N}^k$, zdefiniujmy $P^{\oplus} = \{ p_1 + p_2 + \ldots + p_m \ : \ m \geq 0, \ p_1, p_2, \ldots, p_m \in P \}$.

Zbiór *liniowy* $X \subseteq \mathbb{N}^k$ to dowolny zbiór postaci

$$X = b + P^{\oplus} = \{ b + p_1 + \ldots + p_m \ : \ p_1, \ldots, p_m \in P \}, \quad \text{dla } b \in \mathbb{N}^k, \ P \subseteq_{\text{fin}} \mathbb{N}^k.$$
$$= \{ b + p \ : \ b \in B, \ p \in P^{\oplus} \}$$

b baza, P okresy

Zbiór *semiliniowy* $X \subseteq \mathbb{N}^k$ to dowolna skończona suma zbiorów liniowych:

$$b_1 + {P_1}^{\oplus} \ \cup \ b_2 + {P_2}^{\oplus} \ \cup \ \ldots \cup \ b_n + {P_n}^{\oplus}$$

Pytanie

Niech $B, P \subseteq_{\text{fin}} \mathbb{N}^k$. Czy $B + P^{\oplus} = \{ b + p \ : \ b \in B, \ p \in P^{\oplus} \}$ jest (semi)liniowy?

Dla $k = 1$, zbiory semiliniowe to zbiory prawie okresowe.

Twierdzenie

Zbiór nieujemnych całkowitych rozwiązań układu równań liniowych o całkowitych współczynnikach jest postaci $B + P^{\oplus}$.

Obraz przemienny jest semiliniowy

Twierdzenie (Parikh 1961)

Obraz przemienny języka bezkontekstowego jest semiliniowy.

Wniosek

Dla każdego języka bezkontekstowego L istnieje język regularny R t.że

$$\mathcal{P}(L) \;=\; \mathcal{P}(R).$$

Dowód:

$$b_1 + P_1{}^{\oplus} \quad \cup \quad \ldots \; \cup \quad b_n + P_n{}^{\oplus}$$

$$v + \{v_1, \ldots, v_m\}^{\oplus} \subseteq \mathbb{N}^k \qquad \longmapsto \qquad w\;(w_1 + \ldots + w_m)^* \subseteq A^*$$

Wniosek

Dla alfabetów jednoliterowych, (języki bezkontekstowe) $=$ (języki regularne).

Obraz przemienny jest semiliniowy (dowód)

Niech $L = L(\mathcal{G})$, gdzie $\mathcal{G} = (A, N, S, \alpha)$. Pokażemy, że $\mathcal{P}(L)$ jest semiliniowy.

- Dla $M \subseteq N$, niech $L_M \subseteq L$ zawiera słowa posiadające *M-drzewo wyprowadzenia*, tz. takie, w którym pojawiają się wszystkie nieterminale z M (i żadne inne).

- Ponieważ
$$L = \bigcup_{M \subseteq N} L_M,$$
wystarczy pokazać, że $\mathcal{P}(L_M)$ jest semiliniowy, dla dowolnego $M \subseteq N$.

- Drzewo wyprowadzenia nazwijmy *płytkim*, jeśli na żadnej gałęzi żaden nieterminal nie pojawia się więcej niż $|N|$ razy.

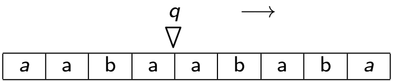
- Pokażemy, że $\mathcal{P}(L_M) = B_M + P_M{}^{\oplus}$, gdzie
$$B_M = \{\, \mathcal{P}(w) \ : w \text{ ma } \textit{płytkie M-drzewo wyprowadzenia} \,\}$$
$$P_M = \{\, \mathcal{P}(wv) \ : wXv \text{ ma } \textit{płytkie } M'\textit{-drzewo wyprowadzenia} \text{ o korzeniu } X, \text{ dla pewnego } X \in N,\ M' \subseteq M \,\}.$$

rysunek

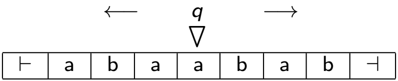
Plan

Dwukierunkowość

Automaty jednokierunkowe:



Automaty dwukierunkowe:



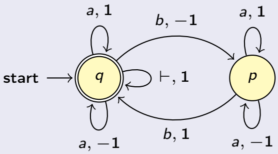
Automat dwukierunkowy

(Niedeterministyczny) *automat dwukierunkowy* $\mathcal{A} = (A, \vdash, \dashv, Q, I, F, \delta)$

$$\delta \subseteq Q \times (A \cup \{\vdash, \dashv\}) \times Q \times \{-1, 0, 1\}$$

$(q, a, q', k) \in \delta$: czytaj a, zmień stan z q na q', zmień pozycję o k

Przykład



($\dashv$ nieużywane)

Konfiguracje automatu dwukierunkowego

Ustalmy automat dwukierunkowy $\mathcal{A} = (A, \vdash, \dashv, Q, I, F, \delta)$ i słowo

$$w = a_1 \ldots a_n \in A^*.$$

Konfiguracja automatu $\mathcal{A}$ na słowie w to para $(q, i) \in Q \times \{0 \ldots n+1\}$

$$\longleftarrow \qquad q \qquad \longrightarrow$$

$$\nabla$$

$\vdash$	a	b	a	a	b	a	b	$\dashv$
0	1		...	i			n	n+1

Konfiguracje początkowe: $I \times \{1\}$
Konfiguracje akceptujące: $F \times \{n+1\}$

Zabraniamy przejść postaci

$$(q, \vdash, q', -1) \qquad (q, \dashv, q', 1),$$

czyli: $\delta \cap (Q \times \{\vdash\} \times Q \times \{-1\} \cup Q \times \{\dashv\} \times Q \times \{1\}) = \emptyset$

Język automatu dwukierunkowego

Definiujemy relację przejścia pomiędzy *konfiguracjami* automatu $\mathcal{A}$ na słowie w.

$$(q, i) \longrightarrow (q', i + k)$$

wtw. gdy

- $1 \leq i \leq n$ oraz δ zawiera przejście (q, a_i, q', k), lub
- $i = 0$ oraz δ zawiera przejście $(q, \vdash, q', k)$, lub
- $i = n + 1$ oraz δ zawiera przejście $(q, \dashv, q', k)$.

Bieg na słowie w to ciąg konfiguracji

$$(q_0, i_0) \longrightarrow (q_1, i_1) \longrightarrow \ldots \longrightarrow (q_m, i_m)$$

gdzie $q_0 \in I$ oraz $i_0 = 1$. Bieg jest akceptujący jeśli $q_m \in F$ oraz $i_m = n + 1$.

Pytanie

Jak długi może być bieg automatu dwukierunkowego na słowie w ?

Język rozpoznawany przez $\mathcal{A}$:

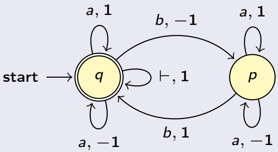
$$L(\mathcal{A}) \;=\; \{\, w \in A^* \;:\; \mathcal{A} \text{ ma bieg akceptujący na } w \,\}.$$

Język automatu dwukierunkowego

$$L(\mathcal{A}) \;=\; \{\, w \in A^* \;:\; \mathcal{A} \text{ ma bieg akceptujący na } w \,\}.$$

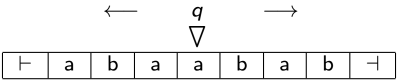
Pytanie

Jaki język rozpoznaje ten automat?



(⊢, ⊣ nieużywane)

Automaty dwukierunkowe jako maszyny Turinga



Automaty dwukierunkowe $=$ maszyny Turinga z taśmą tylko do odczytu

Deterministyczne automaty dwukierunkowe

Automat dwukierunkowy $\mathcal{A} = (A, \vdash, \dashv, Q, I, F, \delta)$ jest *deterministyczny*, jeśli relacja przejścia jest funkcją:

$$\delta : Q \times (A \cup \{\vdash, \dashv\}) \to Q \times \{-1, 0, 1\}$$

Pytanie

Jaki język rozpoznaje ten automat?

- $A = \{a, b\}$
- $Q = \{q_0, q_1, q_2, p_1, p_2, a\}$
- $I = \{q_0\}$
- $F = \{a\}$

	$\vdash$	a	b	$\dashv$
q_0	$(q_0, +1)$	$(q_1, +1)$	$(q_0, +1)$	$(p_0, -1)$
q_1		$(q_2, +1)$	$(q_1, +1)$	
q_2		$(q_0, +1)$	$(q_2, +1)$	
p_0	$(a, +1)$	$(p_0, -1)$	$(p_1, -1)$	
p_1		$(p_1, -1)$	$(p_0, -1)$	
a	$(a, +1)$	$(a, +1)$	$(a, +1)$	

Pytanie

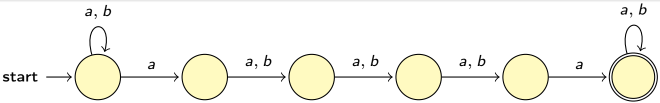
Ile stanów musi mieć deterministyczny automat dwukierunkowy dla języka

$$L_n \;=\; A^* \, a \, A^{n-1}?$$

Pytanie

Ile stanów musi mieć deterministyczny automat dwukierunkowy dla języka

$$L_n \;=\; A^* \, a \, A^{n-1} \, a \, A^*?$$



Odpowiedź

```
idź w prawo do pierwszej a
idź n kroków w prawo
jeśli a to akceptuj
w.p.p.
    idź n − 1 kroków w lewo
    kontynuuj od pierwszej instrukcji

wyjątek:  jeśli ⊣ to odrzuć
```

Automaty dwukierunkowe a jednokierunkowe

Pytanie

Czy automaty dwukierunkowe rozpoznają więcej języków niż jednokierunkowe?

Twierdzenie (Rabin, Scott 1959; Shepardson 1959)

(Automaty dwukierunkowe) $\equiv$ *(automaty jednokierunkowe).*

Dowód (Vardi 1989):

Niech $\mathcal{A} = (A, \vdash, \dashv, Q, I, F, \delta)$ automat dwukierunkowy.

Fakt

$w = a_1 \ldots a_n \notin L(\mathcal{A})$ *wtw. gdy* $\exists P_0, P_1, \ldots, P_{n+1} \subseteq Q$ *t.że*

- $I \subseteq P_1$
- $\forall i \in \{0 \ldots n+1\},\ (q, a_i, q', k) \in \delta.\ q \in P_i \implies q' \in P_{i+k}$
- $F \cap P_{n+1} = \emptyset$ $\qquad (a_0 = \vdash,\ a_{n+1} = \dashv)$

P_0	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
$\vdash$	a	b	a	a	b	a	b	$\dashv$

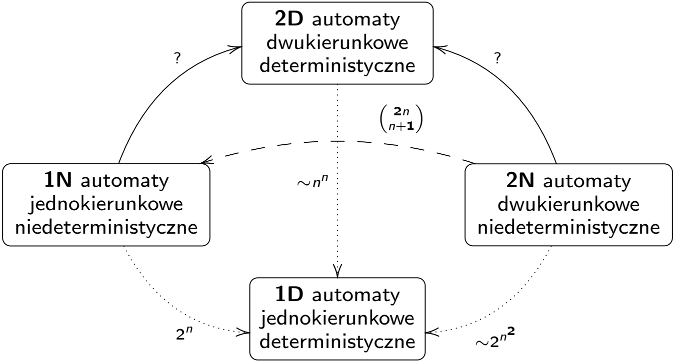
P_0	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
$\vdash$	a	b	a	a	b	a	b	$\dashv$

$$(P_0, P_1) \xrightarrow{a} (P_1, P_2) \xrightarrow{b} \ \ldots\ \xrightarrow{a} (P_6, P_7) \xrightarrow{b} (P_7, P_8)$$

Dowód (c.d.): Definiujemy niedeterministyczny automat jednokierunkowy $\mathcal{A}'$:

- $Q' \ = \ \mathcal{P}(Q) \times \mathcal{P}(Q)$
- $\delta' \ = \ \{((P, P'), a, (P', P'')) \ : \ \begin{array}{l} \forall p' \in P', q \in Q.\ (p', a, q, -1) \in \delta \implies q \in P \\ \forall p' \in P', q \in Q.\ (p', a, q, 0) \in \delta \implies q \in P' \\ \forall p' \in P', q \in Q.\ (p', a, q, 1) \in \delta \implies q \in P''\} \end{array}$
- $I' \ = \ \{(P, P') \ : \ \begin{array}{l} I \subseteq P' \\ \forall p \in P, q \in Q.\ (p, \vdash, q, 0) \in \delta \implies q \in P \\ \forall p \in P, q \in Q.\ (p, \vdash, q, 1) \in \delta \implies q \in P'\} \end{array}$
- $F' \ = \ \{(P, P') \ : \ \begin{array}{l} P' \cap F = \emptyset \\ \forall p' \in P', q \in Q.\ (p', \dashv, q, 0) \in \delta \implies q \in P' \\ \forall p' \in P', q \in Q.\ (p', \dashv, q, -1) \in \delta \implies q \in P\} \end{array}$

Z faktu z poprzedniego slajdu wynika: $w \in L(\mathcal{A}') \iff w \notin L(\mathcal{A})$


2D automaty
dwukierunkowe
deterministyczne
1N automaty
jednokierunkowe
niedeterministyczne
2N automaty
dwukierunkowe
niedeterministyczne
1D automaty
jednokierunkowe
deterministyczne
?
?
$\binom{2n}{n+1}$
$\sim n^n$
2^n
$\sim 2^{n^2}$

W następnym odcinku: