

Języki, automaty i obliczenia

Wykład 3: Języki regularne a automaty skończone

Sławomir Lasota

Uniwersytet Warszawski

11 marca 2015

- $\delta \subseteq Q \times (A \cup \{\varepsilon\}) \times Q$

- Puste przejścia (ε -przejścia)

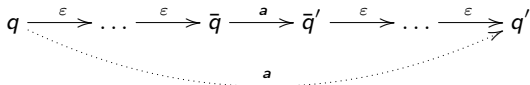
$$\delta(q, \varepsilon, q') \quad q \xrightarrow{\varepsilon} q'$$

- Rozszerzamy puste przejścia:

$$\widehat{\delta}(q, \varepsilon, q') \quad q \xrightarrow{\varepsilon} \xrightarrow{\varepsilon} \dots \xrightarrow{\varepsilon} q'$$

- Usuwanie pustych przejść:

$$(q, a, q') \in \rho \iff \exists \bar{q}, \bar{q}' \cdot \begin{cases} \widehat{\delta}(q, \varepsilon, \bar{q}) \wedge \\ \delta(\bar{q}, a, \bar{q}') \wedge \\ \widehat{\delta}(\bar{q}', \varepsilon, q') \end{cases}$$



Lemat

*Dla każdego wyrażenia regularnego istnieje równoważny automat skończony.
(Każdy język regularny jest rozpoznawany przez automat skończony.)*

Dowód:

Konstruujemy indukcyjnie automat \mathcal{A}_L z pustymi przejściami t. że $L(\mathcal{A}_L) = L$.

- początek:

$$\mathcal{A}_a \ (a \in A) \qquad \mathcal{A}_\varepsilon \qquad \mathcal{A}_\emptyset$$

- krok indukcyjny:

- $\mathcal{A}_L, \mathcal{A}_M \mapsto \mathcal{A}_{LM}$
- $\mathcal{A}_L, \mathcal{A}_M \mapsto \mathcal{A}_{L+M}$
- $\mathcal{A}_L \mapsto \mathcal{A}_{L^*}$

Lemat

*Dla każdego automatu skończonego istnieje równoważne wyrażenie regularne.
(Każdy język rozpoznawany przez automat skończony jest regularny.)*

Dowód:

Bez utraty ogólności załóżmy, że automat ma dokładnie jeden stan początkowy i jeden stan akceptujący:

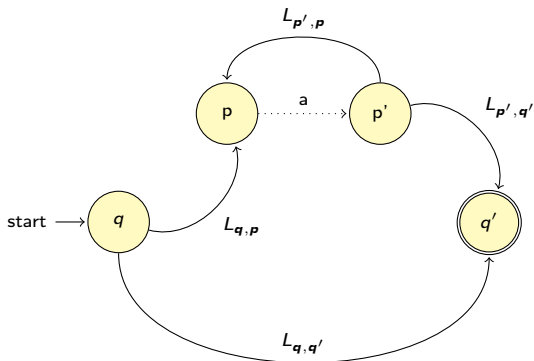
$$L(\mathcal{A}) = \bigcup_{q \in I, q' \in F} L_{qq'}(\mathcal{A})$$

Indukcja ze względu na liczbę przejść $|\delta|$ automatu.

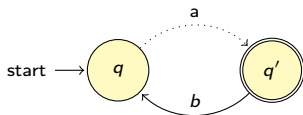
- początek: brak przejść

$$L = \emptyset \quad \text{albo} \quad L = \varepsilon$$

- krok indukcyjny: ...

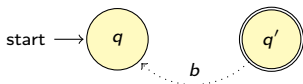


$$L_{q,q'} + L_{q,p} a (L_{p',p} a)^* L_{p',q'}$$



$$L_{q,q'} + L_{q,q} a(L_{q',q} a)^* L_{q',q'} =$$

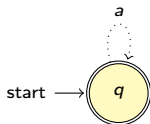
$$\emptyset + \varepsilon a(ba)^* \varepsilon = a(ba)^*$$



$$L_{q',q} = \emptyset + \varepsilon b(\emptyset b)^* \varepsilon = b$$

$$L_{q,q} = \varepsilon$$

$$L_{q,q'} = \emptyset$$

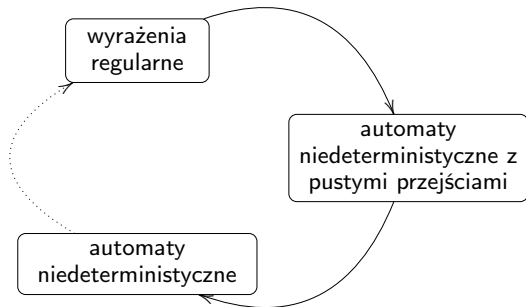


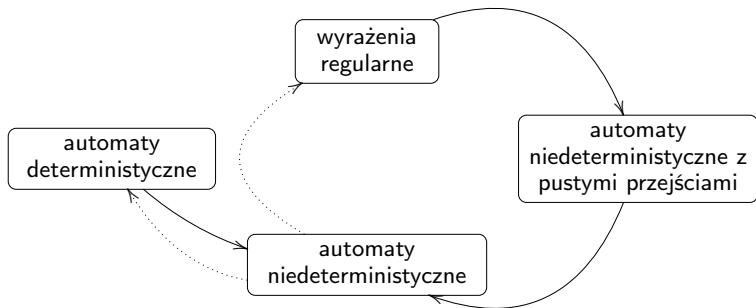
$$\varepsilon + \varepsilon a(\varepsilon a)^* \varepsilon = a^*$$

Twierdzenie (Kleene 1956)

Wyrażenia regularne są równoważne automatom skończonym.

(Automaty skończone rozpoznają dokładnie języki regularne.)





Automat deterministyczny \mathcal{A}_n :

- $A = \{(j, k) \in \{1 \dots n\}^2 : j \neq k\}$
- $Q = \{1 \dots n\} \cup \{\text{śmietnik}\}$
- $I = \{1\}$
- $F = \{n\}$
- $\delta(i, (j, k)) = \begin{cases} k & \text{jeśli } i = j \\ \text{śmietnik} & \text{wpp.} \end{cases}$
 $\delta(\text{śmietnik}, (j, k)) = \text{śmietnik}$

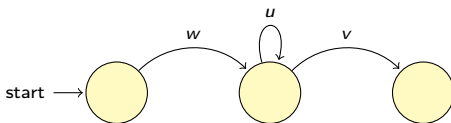
Wyrażenie regularne równoważne automатовi \mathcal{A}_n ma rozmiar $\Omega(2^{n-1})$.

$A = \{a, b\}$. Język $L_n = A^* a A^{n-2}$.

Automat deterministyczny rozpoznający L_n ma $\Omega(2^{n-1})$ stanów.

Obserwacja

Bieg automatu o n stanach, o długości większej niż n , odwiedza dwukrotnie jakiś stan.



Obserwacja

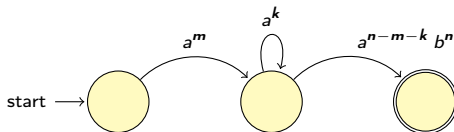
Jeśli automat o n stanach akceptuje słowo o długości większej lub równej n , to słowo to można przedstawić jako

$$w u v, \quad \text{gdzie } |u| > 0 \text{ i } |w u| \leq n$$

i automat akceptuje słowo $w u^m v$, dla każdego $m \in \mathbb{N}$.

Pytanie

$A = \{a, b\}$. Czy język $L = \{a^n b^n : n \in \mathbb{N}\}$ jest regularny?



Odpowiedź

Nie: Dowód nie-wprost.

Założmy, że L jest rozpoznawany przez automat \mathcal{A} o n stanach. \mathcal{A} akceptuje $a^n b^n \in L$, więc $a^{n+k} b^n \in L$ dla pewnego $k > 0$. Sprzeczność.

Lemat o pompowaniu

Dla każdego języka regularnego L istnieje n takie, że każde słowo $x \in L$, $|x| \geq n$, można przedstawić jako

$$x = w u v, \text{ gdzie } |u| > 0 \text{ i } |w u| \leq n$$

i dla każdego $m \in \mathbb{N}$, $w u^m v \in L$.

$$L \text{ regularny} \implies \exists n \forall x \exists x = wuv \forall m. w u^m v \in L$$

Wniosek

L jest nieregularny, jeśli dla każdego n , istnieje słowo $x \in L$, $|x| \geq n$ takie, że dla każdego przedstawienia x jako

$$x = w u v, \text{ gdzie } |u| > 0 \text{ i } |w u| \leq n,$$

istnieje $m \in \mathbb{N}$ takie, że $w u^m v \notin L$.

$$L \text{ nieregularny} \iff \forall n \exists x \forall x = wuv \exists m. w u^m v \notin L$$

L nieregularny $\iff \forall n \exists x \forall x = wuv \exists m. w u^m v \notin L$

Pytanie

$A = \{a, b\}$. Czy język $L = \{a^m b^n : m \leq n\}$ jest regularny?

Odpowiedź

Nie. Dla *dowolnego* $n \in \mathbb{N}$, rozważmy słowo $x_n = a^n b^n \in L$.
Rozważmy *dowolne* słowa w, u, v takie, że

$$x_n = w u v, |u| > 0 \text{ i } |w u| \leq n.$$

Czyli $w \in a^*$, $u \in a^+$ i $v \in a^* b^n$. Wtedy $w u^2 v \notin L$. Zatem L nieregularny.

L nieregularny $\iff \forall n \exists x \forall x = wuv \exists m. w u^m v \notin L$

Pytanie

Czy język palindromów $L = \{w \in \{a, b\}^* : w = w^R\}$ jest regularny?

Odpowiedź

Nie. Dla *dowolnego* $n \in \mathbb{N}$, rozważmy słowo $x_n = a^n b a^n \in L$.
Rozważmy *dowolne* słowa w, u, v takie, że

$$x_n = w u v, |u| > 0 \text{ i } |w u| \leq n.$$

Czyli $w \in a^*$, $u \in a^+$ i $v \in a^* b a^n$. Wtedy $w v \notin L$. Zatem L nieregularny.

Pytanie

Czy język $L = \{(c^* a)^n c^* (b c^*)^n : n \in \mathbb{N}\}$ jest regularny?

Odpowiedź 1

Nie. Dla dowolnego $n \in \mathbb{N}$ rozważmy słowo $x_n = a^n b^n \in L \dots$

Odpowiedź 2

Nie. Dowód nie-wprost.

Przypuśćmy, że L jest regularny. Rozważmy homomorfizm $h : \{a, b, c\}^* \rightarrow \{a, b\}^*$ wyznaczony przez

$$a \mapsto a, b \mapsto b, c \mapsto \varepsilon.$$

Skoro L jest regularny, to $\vec{h}(L) = \{a^n b^n : n \in \mathbb{N}\}$ też. Sprzeczność.

Pytanie

Czy zbiór wyrażeń regularnych L jest językiem regularnym?

Alfabet to $A' = A \cup \{+, *, \varepsilon, \emptyset, (,)\}$.

Pytanie

Czy zbiór wyrażeń regularnych L jest językiem regularnym?

Alfabet to $A' = A \cup \{+, *, \epsilon, \emptyset, (,)\}$.

Odpowiedź

Nie. Dowód nie-wprost.

Przypuśćmy, że L jest regularny. Rozważmy homomorfizm

$$h : (A')^* \rightarrow \{(,)\}^*,$$

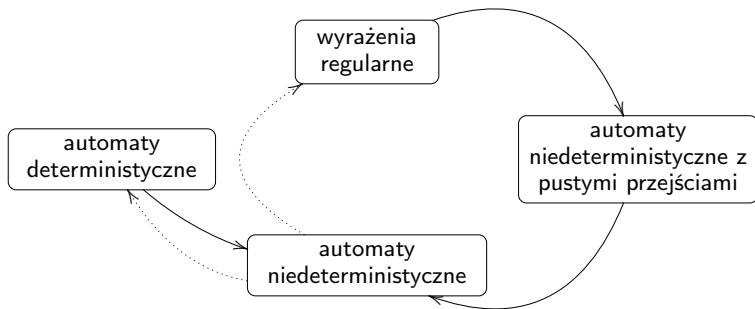
który zachowuje symbole $($ i $)$ a „wymazuje” pozostałe. Zatem $\vec{h}(L)$ to poprawnie zbudowane wyrażenia nawiasowe. Skoro L jest regularny, to $\vec{h}(L)$ też.

Dla dowolnego $n \in \mathbb{N}$, rozważmy słowo $x_n = \underbrace{(\dots(}_{n} \underbrace{) \dots)}_{n} \in L \dots$

Problem decyzyjny

Dane: język regularny L i słowo w

Wynik: czy $w \in L$?



Problem decyzyjny

Dane: automat niedeterministyczny $\mathcal{A} = (A, Q, I, F, \delta)$ i
słowo $w = a_1 \dots a_n$

Wynik: czy $w \in L(\mathcal{A})$?

Algorytm

$X := I$

powtarzaj dla $i = 1 \dots n$

$X := \vec{\delta}(X, a_i) \quad \vec{\delta}(X, a_i) = \{q' \in Q : \exists q \in X. q \xrightarrow{a_i}_{\mathcal{A}} q'\}$

wynik := $(X \cap F \neq \emptyset)$

Problem decyzyjny

Dane: automat niedeterministyczny $\mathcal{A} = (A, Q, I, F, \delta)$

Wynik: czy $L(\mathcal{A}) \neq \emptyset$?

Algorytm

$X := \emptyset$

$Y := I$

dopóki $Y \neq \emptyset$

$X := X \cup Y$

$Y := \vec{\delta}(Y) - X$ $\vec{\delta}(Y) = \{q' \in Q : \exists q \in Y, a \in A. q \xrightarrow{a}_{\mathcal{A}} q'\}$

wynik := $(X \cap F \neq \emptyset)$

Problem decyzyjny

Dane: automat niedeterministyczny $\mathcal{A} = (A, Q, I, F, \delta)$

Wynik: czy $L(\mathcal{A})$ nieskończony?

Problem decyzyjny

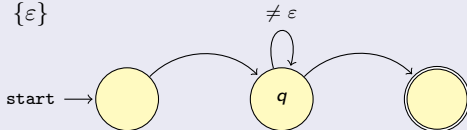
Dane: automat niedeterministyczny $\mathcal{A} = (A, Q, I, F, \delta)$

Wynik: czy $L(\mathcal{A})$ nieskończony?

Algorytm

sprawdź, czy istnieje stan q taki, że

- $L_{I,q}(\mathcal{A}) \neq \emptyset$
- $L_{q,F}(\mathcal{A}) \neq \emptyset$
- $L_{q,q}(\mathcal{A}) \neq \{\varepsilon\}$



Problem decyzyjny

Dane: automaty niedeterministyczne \mathcal{A}, \mathcal{B}

Wynik: czy $L(\mathcal{A}) \subseteq L(\mathcal{B})$?

Problem decyzyjny

Dane: automaty niedeterministyczne \mathcal{A}, \mathcal{B}

Wynik: czy $L(\mathcal{A}) \subseteq L(\mathcal{B})$?

Algorytm

oblicz automat $\bar{\mathcal{B}}$ dla dopełnienia języka $L(\mathcal{B})$

sprawdź, czy $L(\mathcal{A}) \cap L(\bar{\mathcal{B}}) \neq \emptyset$

Problem decyzyjny

Dane: automaty niedeterministyczne \mathcal{A}, \mathcal{B}

Wynik: czy $L(\mathcal{A}) \subseteq L(\mathcal{B})$?

Algorytm

oblicz automat $\bar{\mathcal{B}}$ dla dopełnienia języka $L(\mathcal{B})$

sprawdź, czy $L(\mathcal{A}) \cap L(\bar{\mathcal{B}}) \neq \emptyset$

Problem decyzyjny

Dane: automaty niedeterministyczne \mathcal{A}, \mathcal{B}

Wynik: czy $L(\mathcal{A}) = L(\mathcal{B})$?

W następnym odcinku: minimalizacja automatu

