# Języki, automaty i obliczenia

## Wykład 2: Automaty skończone

Sławomir Lasota

Uniwersytet Warszawski

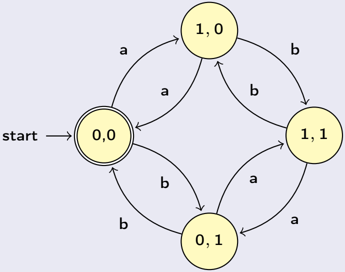
9 marca 2016

# Plan


1. Automaty skończone
2. Determinizacja
3. Operacje na językach

$$L = (aa + bb + (ab + ba)(aa + bb)^*(ab + ba))^*$$

## Przykład



# Automat skończony

$$\mathcal{A} \;=\; (A, Q, I, F, \delta)$$

- $A$ – alfabet
- $Q$ – skończony zbiór stanów
- $I \subseteq Q$ – stany początkowe
- $F \subseteq Q$ – stany akceptujące
- $\delta \subseteq Q \times A \times Q$ – relacja przejścia

## Notacja

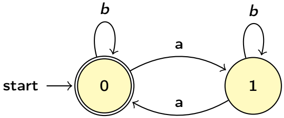
trójkę $(q, a, q') \in \delta$ nazywamy *przejściem*, albo *tranzycją*

zamiast $(q, a, q') \in \delta$ możemy pisać $q \xrightarrow{a} q'$, albo $q \xrightarrow{a}_{\mathcal{A}} q'$

Automat jest *deterministyczny* jeśli

- $\delta$ jest funkcją $Q \times A \to Q$,
- $I$ zawiera jeden stan, $I = \{q_0\}$.

# Przykład: automat deterministyczny

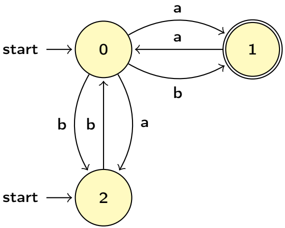



$\mathcal{A} \; = \; (A, Q, I, F, \delta)$

- $A \; = \; \{a, b\}$
- $Q \; = \; \{0, 1\}$
- $I \; = \; F \; = \; \{0\}$
- $\delta \; = \; \{(0, b, 0), (1, b, 1), (0, a, 1), (1, a, 0)\}$

$L \; = ?$

# Przykład: automat niedeterministyczny




$\mathcal{A} = (A, Q, I, F, \delta)$

- $A = \{a, b\}$
- $Q = \{0, 1, 2\}$
- $I = \{0, 2\}$
- $F = \{1\}$
- $\delta = \{(0, a, 1), (0, b, 1), (0, a, 2), (0, b, 2), (1, a, 0), (2, b, 0)\}$

$L = ?$

# Rozszerzona relacja przejścia

Rozszerzamy relację przejścia do relacji $\widehat{\delta} \subseteq Q \times A^* \times Q$:

- $\widehat{\delta}(q, \varepsilon, q)$
- jeśli $\widehat{\delta}(q, w, q')$ i $\delta(q', a, q'')$ to $\widehat{\delta}(q, wa, q'')$

$$\begin{tikzcd} q \arrow[r, dotted, "w"] \arrow[rr, dotted, bend right, "wa"'] & q' \arrow[r, "a"] & q'' \end{tikzcd}$$

**Notacja**

zamiast $(q, w, q') \in \widehat{\delta}$ możemy pisać $q \xrightarrow{w} q'$

Dla automatów deterministycznych: $\widehat{\delta} : Q \times A^* \to Q$.

Stany osiągalne:

$$\{\, q \in Q \;:\; \exists w \in A^*, q_0 \in I.\ q_0 \xrightarrow{w} q \,\}$$

*Bieg* (obliczenie) automatu na słowie $w = a_1 a_2 \ldots a_n$:

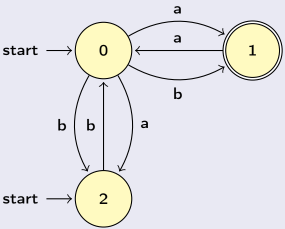
$$q_0 \in I \qquad (q_0, a_1, q_1)\ (q_1, a_2, q_2)\ \ \ldots\ \ (q_{n-1}, a_n, q_n)$$

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \ \ \ldots\ \ q_{n-1} \xrightarrow{a_n} q_n$$

Bieg jest *akceptujący* jeśli $q_n \in F$.

### Pytanie

Ile biegów na słowie *ba* ma ten automat? Czy ma bieg na każdym słowie?




### Fakt

*Automat deterministyczny ma dokładnie jeden bieg na każdym słowie.*

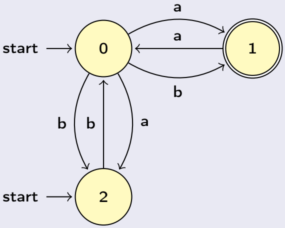
Automat jest *jednoznaczny* jeśli ma co najwyżej jeden bieg akceptujący na każdym słowie.

### Fakt

*Automat deterministyczny ma dokładnie jeden bieg na każdym słowie, więc jest jednoznaczny.*

### Pytanie

Czy ten automat jest jednoznaczny?



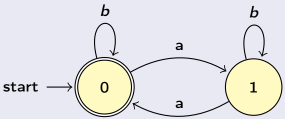
# Język automatu

Język *rozpoznawany* przez automat:

- $L(\mathcal{A}) \stackrel{\text{def}}{=} \{ w \in A^* \ : \ \exists q \in I, q' \in F.\ q \xrightarrow{w} q' \}$
- $L(\mathcal{A}) \stackrel{\text{def}}{=} \{ w \in A^* \ : \ \mathcal{A}$ ma bieg akceptujący na $w \}$
- $L(\mathcal{A}, q) \stackrel{\text{def}}{=} \ \ldots$

**Przykład**

$L(\mathcal{A}) \ = \ b^*(ab^*ab^*)^*$



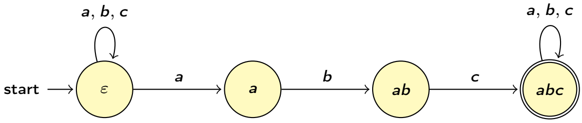

Dla automatów deterministycznych:

$$L(\mathcal{A}) \ = \ \{ w \in A^* \ : \ \widehat{\delta}(q_0, w) \in F \}$$

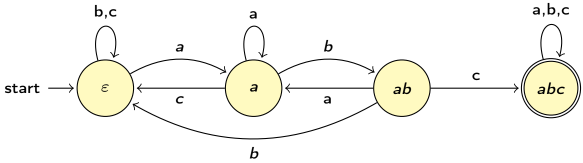
# Przykład – wyszukiwanie wzorca

- $A = \{a, b, c\}$
- $L(\mathcal{A}) = A^* abc\, A^*$

- automat niedeterministyczny:




- automat deterministyczny:



# Przykład – podzielność

$L(\mathcal{A}) \;=\;$ liczby podzielne przez 3

- $A \;=\; \{0, 1, \dots, 9\}$
- $Q \;=\; \{0, 1, 2\}$
- $I \;=\; \{0\}, \quad F = \{0\}$
- $\delta(q, a) \;=\; (q + a) \mod 3$

$L(\mathcal{A}) \;=\;$ (liczby podzielne przez 3) $- \{\varepsilon\}$

- $Q \;=\; \{0, 1, 2, \text{początek}\}$
- $I \;=\; \{\text{początek}\}$
- $\delta(\text{początek}, a) \;=\; a \mod 3$

# Przykład – szachy

Automat $\mathcal{A}$:

- $A =$ wszystkie ruchy w szachach
- $Q \;=\;$ (wszystkie ustawienia figur na planszy szachowej) $\times$ {białe, czarne}
- $I \;=\; \{$(ustawienie początkowe, białe)$\}$
- $F \;=\; \{$ustawienia szach-mat$\}$
- $\delta((u, \text{białe}), r) = (u', \text{czarne})$

  $\delta((u, \text{czarne}), r) = (u', \text{białe}) \qquad u' \;=\; \text{wykonaj}(u, r)$

$L(\mathcal{A}) \;=\;$ wszystkie rozgrywki szachowe zakończone matem

# Plan

1. Automaty skończone
2. Determinizacja
3. Operacje na językach

# Determinizacja?

Równoważność automatów:

$$L(\mathcal{A}) \;=\; L(\mathcal{A}')$$

**Pytanie**

Czy dla każdego automatu skończonego istnieje równoważny automat deterministyczny?

**Twierdzenie**

*Dla każdego automatu skończonego istnieje równoważny automat deterministyczny.*

*Dowód:*

- $Q' \;:=\; \mathcal{P}(Q)$
- $I' \;:=\; \{I\}$
- $F' \;:=\; \{\, X \subseteq Q \;:\; X \cap F \neq \emptyset \,\}$
- $\delta'(X, a) \;:=\; \{\, \bar{q} \in Q \;:\; \exists q \in X.\ \delta(q, a, \bar{q}) \,\}$

Przez indukcję po długości $w$ pokazujemy:

$$\widehat{\delta'}(I, w) \;=\; \{\, \bar{q} \in Q \;:\; \exists q \in I.\ \widehat{\delta}(q, w, \bar{q}) \,\}$$

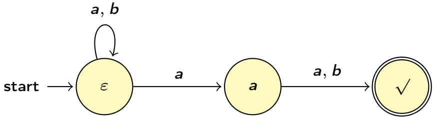
Zatem

$$w \in L(\mathcal{A}') \iff \widehat{\delta'}(I, w) \in F' \iff \widehat{\delta'}(I, w) \;\cap\; F \neq \emptyset \iff$$
$$\iff \exists q \in I, \bar{q} \in F.\ \widehat{\delta}(q, w, \bar{q}) \iff w \in L(\mathcal{A})$$

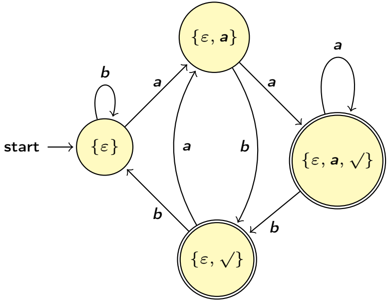
# Przykład

$$A = \{a, b\}$$
$$L = A^* a A$$

- automat niedeterministyczny (ale jednoznaczny):



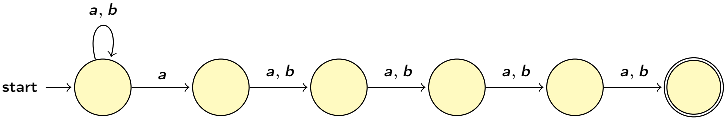

- automat deterministyczny:



# Determinizacja jest wykładnicza

$$\begin{aligned} A &= \{a, b\} \\ L_n &= A^* a A^{n-2} \qquad n \geq 2 \end{aligned}$$

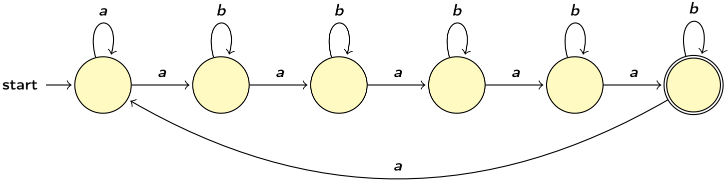
Automat niedeterministyczny (ale jednoznaczny):




**Pytanie**

Ile stanów ma automat deterministyczny?

# Determinizacja jest wykładnicza




## Pytanie

Ile stanów ma automat deterministyczny?

# Plan


1. Automaty skończone

2. Determinizacja

3. Operacje na językach

# Operacje na językach: suma, przecięcie

$$\mathcal{A},\ \mathcal{A}' \ \mapsto\ \mathcal{A}'' \qquad\qquad L(\mathcal{A}'')\ =\ L(\mathcal{A}) \cup L(\mathcal{A}')$$

- $Q'' \ := \ Q \uplus Q'$
- $I'' \ := \ I \uplus I'$
- $F'' \ := \ F \uplus F'$
- $\delta'' \ := \ \delta \uplus \delta'$

$$\mathcal{A},\ \mathcal{A}' \ \mapsto\ \mathcal{A}'' \qquad\qquad L(\mathcal{A}'')\ =\ L(\mathcal{A}) \cap L(\mathcal{A}')$$

- $Q'' \ := \ Q \times Q'$
- $I'' \ := \ I \times I'$
- $F'' \ := \ F \times F'$
- $\delta'' \ := \ \{\, ((q, q'), a, (p, p')) \ : \ (q, a, p) \in \delta,\ (q', a, p') \in \delta' \,\}$

$$(q, q') \xrightarrow{a}_{\mathcal{A}''} (p, p') \quad \text{wtw. gdy} \quad q \xrightarrow{a}_{\mathcal{A}} p \ \text{ i } \ q' \xrightarrow{a}_{\mathcal{A}'} p'$$

# Operacje na językach: dopełnienie

$$\mathcal{A} \mapsto \mathcal{B} \qquad L(\mathcal{B}) = A^* - L(\mathcal{A})$$

**Pytanie**

Czy wystarczy zamienić stany akceptujące z nieakceptującymi?

$$\text{start} \longrightarrow \varepsilon \;(\circlearrowleft a, b) \xrightarrow{a} a \xrightarrow{a,\, b} \checkmark$$

Krok 1: determinizacja $\mathcal{A} \mapsto \mathcal{A}'$

Krok 2: zamiana stanów akceptujących z nieakceptującymi $\mathcal{A}' \mapsto \mathcal{B}$

- $Q_{\mathcal{B}} := Q_{\mathcal{A}'}$
- $I_{\mathcal{B}} := I_{\mathcal{A}'}$
- $F_{\mathcal{B}} := Q_{\mathcal{A}'} - F_{\mathcal{A}'}$
- $\delta_{\mathcal{B}} := \delta_{\mathcal{A}'}$

# Operacje na językach: przeplot

**Przykład**

las $\otimes$ so $=$ {lasso, lsaso, slaso, lasos, lsaos, slaos, lsoas, sloas, solas}

$$w \otimes v \stackrel{\text{def}}{=} \{\, u \in A^* \;:\; \exists X \subseteq \{1 \ldots |u|\}.\ u|_X = w,\, u|_{\{1\ldots|u|\}-X} = v \,\}$$

$$L \otimes M \stackrel{\text{def}}{=} \bigcup_{w \in L, v \in M} w \otimes v$$

$$\mathcal{A},\ \mathcal{A}' \ \mapsto\ \mathcal{A}'' \qquad\qquad L(\mathcal{A}'') \ = \ L(\mathcal{A}) \otimes L(\mathcal{A}')$$

- $Q'' \ := \ Q \times Q'$
- $I'' \ := \ I \times I'$
- $F'' \ := \ F \times F'$

$$(q, q') \xrightarrow{a}_{\mathcal{A}''} (p, p') \quad \text{wtw. gdy} \quad \begin{array}{ll} q \xrightarrow{a}_{\mathcal{A}} p \ \text{ i } \ q' = p', & \text{albo} \\ q' \xrightarrow{a}_{\mathcal{A}'} p' \ \text{ i } \ q = p & \end{array}$$

# W następnym odcinku: wyrażenia regularne = automaty skończone

$$L \;=\; (aa + bb + (ab + ba)(aa + bb)^*(ab + ba))^*$$