Języki, automaty i obliczenia

Wykład 6: Automaty na drzewach
(wykład primaaprilisowy)

Sławomir Lasota

Uniwersytet Warszawski

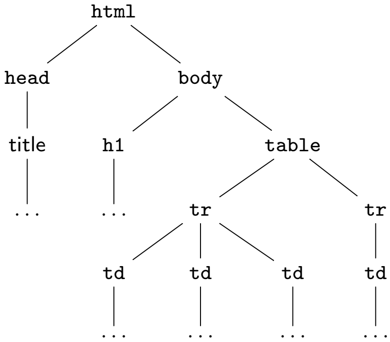
6 kwietnia 2016

Drzewa?

Pytanie

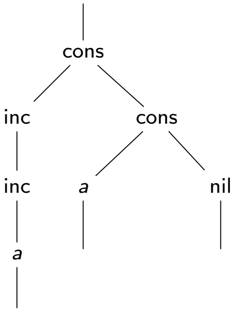
Dlaczego akurat drzewa?

```
<html>
    <head>
        <title> ...</title>
    </head>
    <body>
        <h1> ...</h1>
        <table>
            <tr>
                <td> ...</td>
                <td> ...</td>
                <td> ...</td>
            </tr>
            <tr>
                <td> ...</td>
            </tr>
        </table>
    </body>
</html>
```



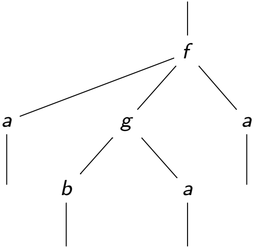
Drzewa – przykłady

$$A = \{a, \mathsf{nil}, \mathsf{inc}, \mathsf{cons}\}$$



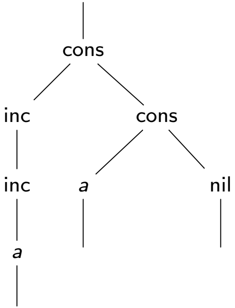
$$\begin{aligned}
\text{arność}(a) &= 1\\
\text{arność}(\mathsf{nil}) &= 1\\
\text{arność}(\mathsf{inc}) &= 1\\
\text{arność}(\mathsf{cons}) &= 2
\end{aligned}$$

$$A = \{a, b, f, g\}$$

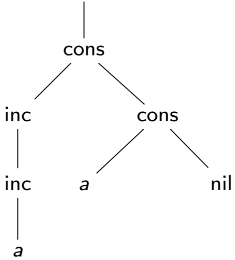


$$\begin{aligned}
\text{arność}(a) &= 1\\
\text{arność}(b) &= 1\\
\text{arność}(f) &= 3\\
\text{arność}(g) &= 2
\end{aligned}$$

Drzewa a termy



$\text{cons}(\text{inc}(\text{inc}(a(\_))), \text{cons}(a(\_), \text{nil}(\_)))$



$\text{cons}(\text{inc}(\text{inc}(a), \text{cons}(a, \text{nil}))$

Pytanie

Jak narysować puste drzewo?

Alfabet A z arnościami:

$$\text{arność} : A \to \mathbb{N} \setminus \{0\}$$

Drzewo nad A to para $t = (T, l)$, gdzie

$$T \subseteq \mathbb{N}^*, \qquad l : T \to A$$

t.że

- T jest zamknięty na prefiksy: $wv \in T \implies w \in T$
- jeśli $w \in T$ i $\text{arność}(l(w)) = n$, to

$$\{\, i \;:\; wi \in T \,\} \;=\; \{0 \ldots n-1\} \qquad \text{albo} \qquad \{\, i \;:\; wi \in T \,\} \;=\; \emptyset$$

Notacja

W szczególnym przypadku, gdy

$$\text{arność}(a) \;=\; 1 \quad \text{dla wszystkich } a \in A,$$

otrzymujemy słowa nad A.

Bieg na słowie

słowo $a_1 a_2 \ldots a_n$:

a_1

a_2

$\ldots$

a_n

bieg akceptujący:

$\boxed{q_0}\ \in I$

a_1

$\boxed{q_1}$

a_2

$\boxed{q_2}$

$\ldots$

$\boxed{q_{n-1}}$

a_n

$\boxed{q_n}\ \in F$

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

$$\delta \subseteq Q \times A \times Q^*$$

t.że

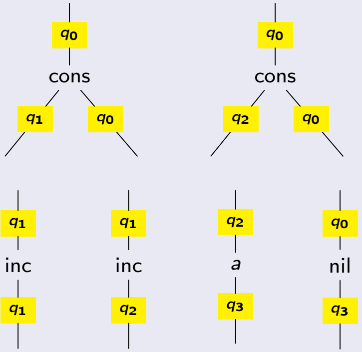
$$(q, a, w) \in \delta \implies |w| = \text{arność}(a)$$

Przykład

- $A = \{a, \text{nil}, \text{inc}, \text{cons}\}$

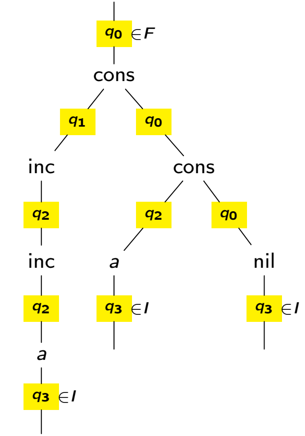
$$\begin{aligned} \text{arność}(a) &= 1 \\ \text{arność}(\text{nil}) &= 1 \\ \text{arność}(\text{inc}) &= 1 \\ \text{arność}(\text{cons}) &= 2 \end{aligned}$$

- $Q = \{q_0, q_1, q_2, q_3\}$
- $I = \{q_0\}$
- $F = \{q_3\}$
- $\delta = \{(q_0, \text{cons}, q_1 q_0), (q_1, \text{inc}, q_1),$
  $\{(q_0, \text{cons}, q_2 q_0), \ldots\}$

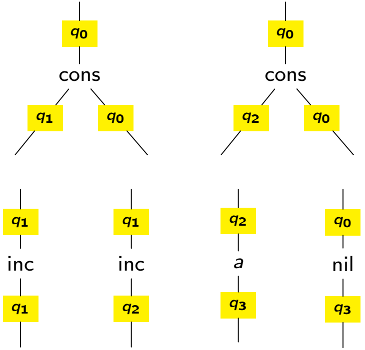


Bieg na drzewie – przykład

bieg akceptujący automatu:



przejścia automatu:



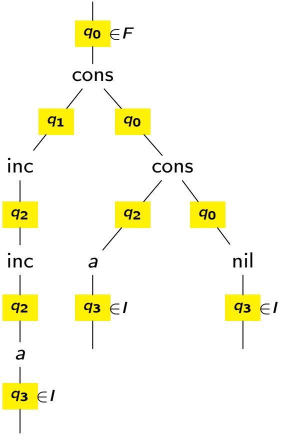
Co tu jest nie tak?

Bieg akceptujący automatu $\mathcal{A}$ na drzewie $t = (T, l)$ nad A to funkcja

$$f : T' \to Q, \qquad \text{gdzie } T' = T \cup \{\, wi \;:\; w \in T,\ 0 \leq i < \text{arność}(l(w))\,\}$$

t.że

- $(f(w), l(w), (f(wi))_{0 \leq i < \text{arność}(l(w))}) \in \delta$, dla każdego $w \in T$
- $f(\varepsilon) \in F$
- $f(w) \in I$, dla każdego $w \in T' - T$



Język $L(\mathcal{A}) = \ldots$ Języki regularne drzew.

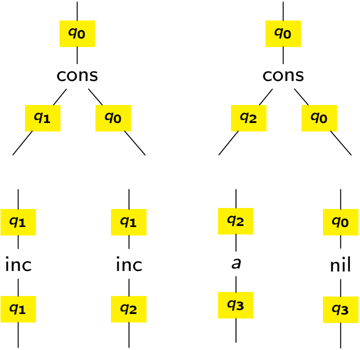
Deterministyczne automaty na drzewach

Automaty deterministyczne „z góry w dół”:

$$\forall q \in Q, a \in A.\ \exists! w \in Q^*.\ (q, a, w) \in \delta \qquad \delta_a : Q \to Q^{\text{arność}(a)}$$

Automaty deterministyczne „z dołu do góry”:

$$\forall a \in A, w \in Q^{\text{arność}(a)}.\ \exists! q \in Q.\ (q, a, w) \in \delta \qquad \delta_a : Q^{\text{arność}(a)} \to Q$$

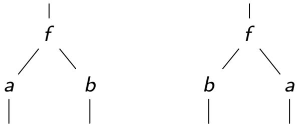


Determinizacja

Pytanie

Czy zachodzi determinizacja „z góry w dół” ?

Brak determinizacji „z góry w dół":



Pytanie

Czy zachodzi determinizacja „z dołu do góry” ?

Tak. Konstrukcja podzbiorowa.

Języki regularne drzew

Pytanie

Które z poniższych języków są regularnymi językami drzew?

Które z nich są deterministyczne „z góry w dół" ?

$$A \;=\; \{a, b\}, \quad \text{arność}(a) \;=\; \text{arność}(b) \;=\; 2$$

- b występuje dokładnie raz
- b nie występuje ani raz
- na każdej ścieżce parzysta liczba a
- na pewnej ścieżce parzysta liczba a
- pod każdym a jest b

$$A \;=\; \{g, f, a, b, c\}, \quad \begin{array}{l} \text{arność}(g) \;=\; 3, \; \text{arność}(f) \;=\; 2 \\ \text{arność}(a) \;=\; \text{arność}(b) \;=\; \text{arność}(c) \;=\; 1 \end{array}$$

- $\{\, f(a^n, b^n) \;:\; n \geq 1 \,\}$
- $\{\, g(c_1, g(c_2, g(\ldots g(c_n, c, c_n), \ldots), c_2), c_1) \;:\; c_1, \ldots, c_n \in \{a, b\} \,\}$

Twierdzenie Myhilla-Nerode'a dla słów

Język L słów nad A określa relację równoważności w zbiorze A^*:

$$w \sim_L v \iff \forall u \in A^*.\ (w\,u \in L \iff v\,u \in L)$$

Twierdzenie (Myhill-Nerode 1958)

Niech $L \subseteq A^*$. Następujące warunki są równoważne:

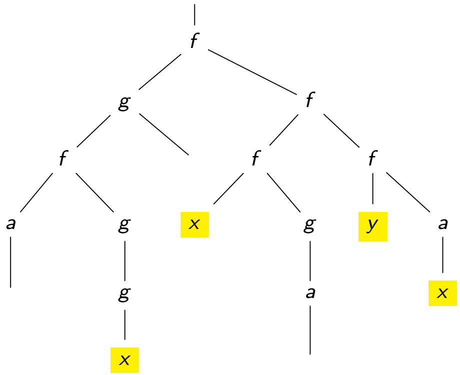
- *L jest regularny*
- *$\sim_L$ ma skończony indeks*

Pytanie

Czym jest „sufiks'' drzewa? Czym jest „prefiks'' drzewa? Czym jest konkatenacja drzew?

Konkatenacja $w\,u$ jako podstawienie w do kontekstu $\_\,u$.

Kontekst $C[x, y]$

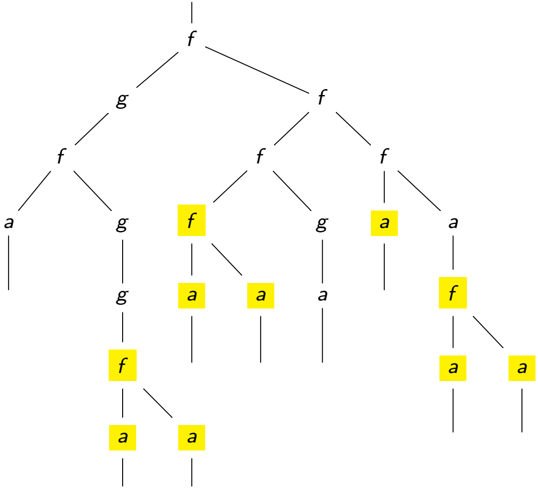


Kontekst trywialny x

Co tu jest nie tak?

Kontekst jest *liniowy* jeśli każda zmienna występuje dokładnie raz.

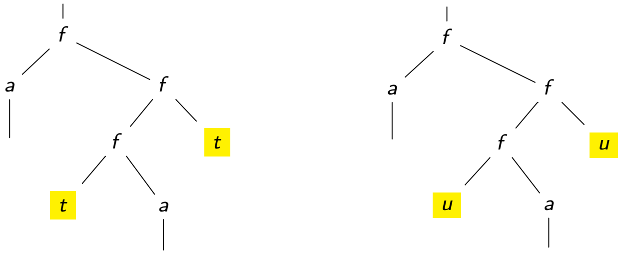
$C[f(a,a),a]$



Równoważność Myhilla-Nerode'a

Język L drzew nad A określa relację równoważności w zbiorze drzew nad A:

$$t \sim_L u \iff \forall C[x].\ (C[t] \in L \iff C[u] \in L)$$



Pytanie

Czy można ograniczyć się do kontekstów liniowych $C[x]$?

Twierdzenie Myhilla-Nerode'a

Język L drzew nad A określa relację równoważności w zbiorze drzew nad A:

$$t \sim_L u \iff \forall C[x].\ (C[t] \in L \iff C[u] \in L)$$

Twierdzenie

Dla języka drzew L nad A, następujące warunki są równoważne:

- *L jest regularny*
- *$\sim_L$ ma skończony indeks*

Minimalizacja podobnie jak dla słów.

Lemat o pompowaniu

Lemat o pompowaniu

Dla każdego regularnego języka drzew L istnieje n takie, że każde drzewo $t \in L$ o głębokości przynajmniej n można przedstawić jako

$$t = C[D[u]],$$

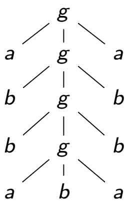
dla pewnych kontekstów liniowych $C[x], D[x]$ i dla pewnego drzewa u, takich, że

- kontekst $D[x]$ jest nietrywialny
- głębokość zmiennej w kontekście $C[D[X]]$ wynosi co najwyżej n,
- dla każdego $m \in \mathbb{N}$, $C[D^m[u]] \in L$.

rysunek: pompowanie drzew

$$A = \{g, a, b\}$$

$$\begin{aligned} \text{arność}(g) &= 3 \\ \text{arność}(a) &= 1 \\ \text{arność}(b) &= 1 \end{aligned}$$



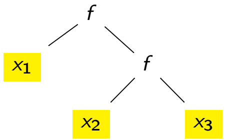
Pytanie

Czym jest homomorfizm z drzew nad A do drzew nad B ?

$$A = \{g, a, b\} \qquad B = \{f, a, b\}$$

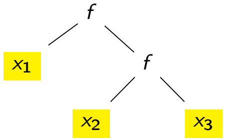
$$\begin{aligned} \text{arność}(g) &= 3 & \qquad \text{arność}(f) &= 2 \\ \text{arność}(a) &= 1 & \qquad \text{arność}(a) &= 1 \\ \text{arność}(b) &= 1 & \qquad \text{arność}(b) &= 1 \end{aligned}$$

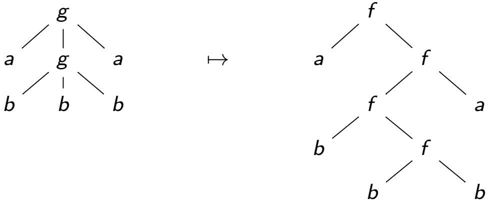
$a \mapsto a \qquad b \mapsto b \qquad g \quad \mapsto$



Homomorfizmy

$a \mapsto a$ $\quad b \mapsto b$ $\quad g \mapsto$





Operacje zachowujące regularność

Języki regularne drzew są zamknięte na następujące operacje:

Co tu jest nie tak?

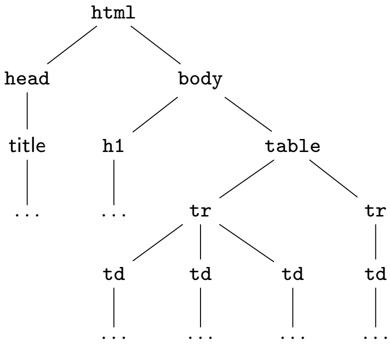
- operacje boolowskie,
- obrazy homomorficzne,
- przeciwobrazy homomorficzne

$$a \mapsto a \qquad\qquad g \;\mapsto\; \begin{array}{ccc} & f & \\ \diagup & & \diagdown \\ \boxed{x} & & \boxed{x} \end{array}$$

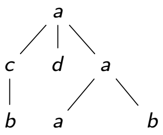
Pytanie

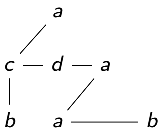
Czy należy ograniczyć się do homomorfizmów liniowych ?

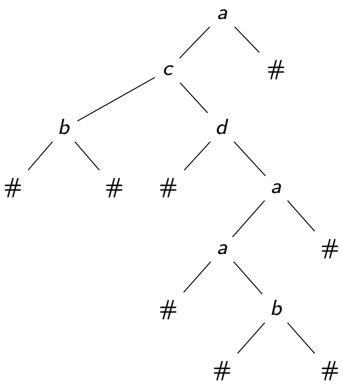
```
<html>
    <head>
        <title> ...</title>
    </head>
    <body>
        <h1> ...</h1>
        <table>
            <tr>
                <td> ...</td>
                <td> ...</td>
                <td> ...</td>
            </tr>
            <tr>
                <td> ...</td>
            </tr>
        </table>
    </body>
</html>
```



Można ograniczyć stopień?



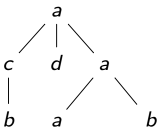


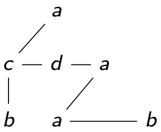


Automaty dla drzew nieograniczonego stopnia

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

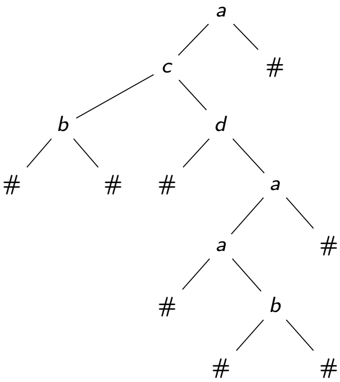
$\delta \subseteq_{\text{fin}} Q \times A \times (\text{języki regularne nad } Q)$.





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

$\delta \subseteq_{\text{fin}} Q \times A \times (Q^2 \cup Q)$.



W następnym odcinku:

języki bezkontekstowe