

Języki, automaty i obliczenia

Wykład 5: Języki bezkontekstowe

Sławomir Lasota

Uniwersytet Warszawski

25 marca 2026

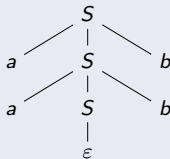
- 1 Gramatyki bezkontekstowe - po co?
- 2 Gramatyki bezkontekstowe
- 3 Języki bezkontekstowe a drzewa
- 4 Postaci normalne i problemy decyzyjne

Przykład

$$L = \{ a^n b^n : n \in \mathbb{N} \}$$

$$S \longrightarrow a S b \mid \varepsilon$$

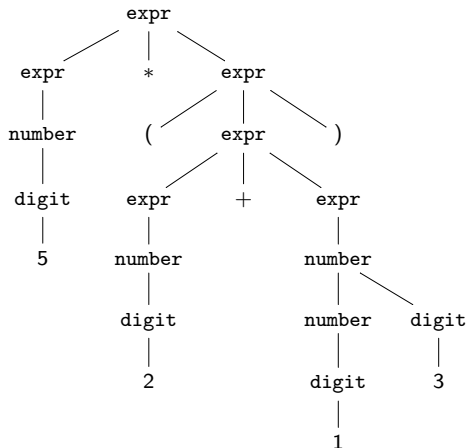
$$S \longrightarrow a S b \longrightarrow a a S b b \longrightarrow a a b b$$



$\text{expr} \rightarrow \text{expr} + \text{expr} \mid \text{expr} * \text{expr} \mid (\text{expr}) \mid \text{number}$
 $\text{number} \rightarrow \text{digit} \mid \text{number digit}$
 $\text{digit} \rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

Drzewo składniowe wyrażenia

5 * (2 + 13)



$$\begin{aligned}\text{expr} &\longrightarrow \text{expr} + \text{expr} \mid \text{expr} * \text{expr} \mid (\text{expr}) \mid \text{number} \\ \text{number} &\longrightarrow \text{digit} \mid \text{number digit} \\ \text{digit} &\longrightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9\end{aligned}$$

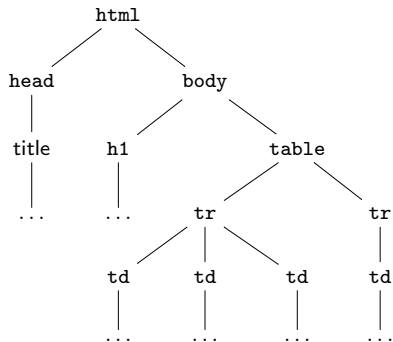
- Analizator składniowy - gramatyka bezkontekstowa:

```
expr :  
    expr '+' expr    { $$ := $1 + $3; }  
    | expr '*' expr  { $$ := $1 * $3; }  
    | '(' expr ')'    { $$ := $2; }  
    | number          { $$ := $1; };
```

- Analizator leksykalny - wyrażenie regularne:

```
number :  
    digit            { $$ := $1; }  
    | number digit   { $$ := 10 * $1 + $2; }
```

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



- 1 Gramatyki bezkontekstowe - po co?
- 2 Gramatyki bezkontekstowe**
- 3 Języki bezkontekstowe a drzewa
- 4 Postaci normalne i problemy decyzyjne

$$\mathcal{G} = (A, N, S, \alpha)$$

- A to skończony zbiór symboli końcowych (*terminalnych*)
- N to skończony zbiór symboli niekończących (*nieterminalnych*), $A \cap N = \emptyset$
- $S \in N$ to symbol początkowy
- $\alpha \subseteq N \times (A \cup N)^*$ to skończony zbiór reguł przepisywania (*produkcji*)

Notacja

Reguły $(X, v) \in \alpha$ będziemy zapisywać $X \rightarrow_{\mathcal{G}} v$ albo $X \rightarrow v$.

Przykład

Gramatyka $\mathcal{G} = (A, N, S, \alpha)$:

- $A = \{a, b, c\}$
- $N = \{S, P\}$
- $\alpha = \{ (S, SS), (S, P), (P, aPa), (P, bPb), (P, \varepsilon) \}$

$$S \rightarrow SS \mid P$$

$$P \rightarrow aPa \mid bPb \mid \varepsilon$$

Reguły rozszerzamy do relacji $\rightarrow_{\mathcal{G}} \subseteq (A \cup N)^* \times (A \cup N)^*$:

$$w \rightarrow_{\mathcal{G}} w' \iff \exists (X \rightarrow_{\mathcal{G}} v) \in \alpha, \exists u, t \in (A \cup N)^*. w = u X t, w' = u v t$$

i domykamy zwrotno-tranzytywnie: $w \twoheadrightarrow_{\mathcal{G}} w'$ wtw. gdy istnieje ciąg

$$w = v_0 \rightarrow_{\mathcal{G}} v_1 \rightarrow_{\mathcal{G}} \dots \rightarrow_{\mathcal{G}} v_n = w', \quad n \geq 0$$

zwany **wyprowadzeniem** (wywodem) w' z w (w gramatyce $\mathcal{G}$).

Język **generowany** przez gramatykę: $L(\mathcal{G}) = \{ w \in A^* : S \twoheadrightarrow_{\mathcal{G}} w \}.$

Przykład

$$S \rightarrow SS \mid P$$

$$P \rightarrow aPa \mid bPb \mid \varepsilon$$

Słowo $ab b a b b \in L(\mathcal{G})$, ponieważ ma wyprowadzenie:

$$\begin{aligned} S &\rightarrow_{\mathcal{G}} SS \rightarrow_{\mathcal{G}} PS \rightarrow_{\mathcal{G}} aPaS \rightarrow_{\mathcal{G}} a bPbaS \rightarrow_{\mathcal{G}} a bbaS \rightarrow_{\mathcal{G}} \\ &abbaP \rightarrow_{\mathcal{G}} ab b a b P b \rightarrow_{\mathcal{G}} ab b a b b \end{aligned}$$

$$L(\mathcal{G}) = ?$$

Przykład

$$S \longrightarrow SS \mid P$$

$$P \longrightarrow aPa \mid bPb \mid \varepsilon$$

$$L(\mathcal{G}) = (\text{palindromy parzyste})^*$$

$$S \longrightarrow SS \mid P$$

$$P \longrightarrow aSa \mid bSb \mid \varepsilon$$

$$L(\mathcal{G}) = ?$$

$$S \longrightarrow SS \mid P$$

$$P \longrightarrow aPa \mid bPb \mid \varepsilon$$

Pytanie

Czy każde słowo ma co najwyżej jedno wyprowadzenie w gramatyce?

Nie, np.

$$\begin{aligned} S &\longrightarrow_G SS \longrightarrow_G PS \longrightarrow_G aPaS \longrightarrow_G abPbaS \longrightarrow_G abbaS \longrightarrow_G \\ &abbaP \longrightarrow_G abbabPb \longrightarrow_G abbabb \end{aligned}$$

$$\begin{aligned} S &\longrightarrow_G SS \longrightarrow_G SP \longrightarrow_G SbPb \longrightarrow_G Sbb \longrightarrow_G \\ Pbb &\longrightarrow_G aPabb \longrightarrow_G abPbabb \longrightarrow_G abbabb \end{aligned}$$

Przykłady języków bezkontekstowych

- $\{a^n b^n : n \in \mathbb{N}\}$

$$S \longrightarrow aSb \mid \varepsilon$$

- palindromy nad $\{a, b\}$:

$$P \longrightarrow aPa \mid bPb \mid a \mid b \mid \varepsilon$$

- $\{w \in \{a, b\}^* : \#_w(a) = \#_w(b)\}$

$$X \longrightarrow aXb \mid bXa \mid aXbbXa \mid bXaaXb \mid \varepsilon$$

- wyrażenia arytmetyczne:

$$W \longrightarrow W \cdot W \mid W + W \mid (W) \mid 0 \mid 1$$

$$S \longrightarrow S + I \mid I$$

$$I \longrightarrow I \cdot C \mid C$$

$$C \longrightarrow (S) \mid 0 \mid 1$$

- poprawnie zbudowane wyrażenia nawiasowe:

$$W \longrightarrow WW \mid [W] \mid (W) \mid \varepsilon$$

Fakt

Każdy język regularny jest bezkontekstowy.

Dowód:

Niech $\mathcal{A} = (A, Q, \{q_0\}, F, \delta)$. Definiujemy gramatykę $\mathcal{G} = (A, N, S, \alpha)$:

- $N = Q$
- $S = q_0$
- $\alpha = \{q \rightarrow a q' : q \xrightarrow{a}_{\mathcal{A}} q'\} \cup \{q \rightarrow \varepsilon : q \in F\}$

$$w \in L(\mathcal{G}) \iff q_0 \twoheadrightarrow_{\mathcal{G}} w \iff \exists q \in F. \hat{\delta}(q_0, w, q) \iff w \in L(\mathcal{A})$$

Wyprowadzenie:

$$w \rightarrow_G w' \iff \exists (X \rightarrow_G v) \in \alpha, \quad u, t \in (A \cup N)^* : w = uXt, \quad w' = uv t$$

Wyprowadzenie *lewostronne*:

$$w \xrightarrow{L}_G w' \iff \exists (X \rightarrow_G v) \in \alpha, \quad u \in A^*, t \in (A \cup N)^* : w = uXt, \quad w' = uv t$$

Przykład

$$\begin{aligned} S &\xrightarrow{L}_G SS & SS &\xrightarrow{L}_G PS & PS &\xrightarrow{L}_G aPaS & aPaS &\xrightarrow{L}_G abPbaS & abPbaS &\xrightarrow{L}_G abb aS & abb aS &\xrightarrow{L}_G \\ &abb aP & abb aP &\xrightarrow{L}_G abb abPb & abb abPb &\xrightarrow{L}_G abb abbb \end{aligned}$$

Fakt

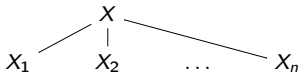
Jeśli $w \twoheadrightarrow_G w' \in A^*$ to $w \xrightarrow{L}_G w'$.

Wyprowadzenie *prawostronne*?

- 1 Gramatyki bezkontekstowe - po co?
- 2 Gramatyki bezkontekstowe
- 3 Języki bezkontekstowe a drzewa**
- 4 Postaci normalne i problemy decyzyjne

Drzewo wyprowadzenia w gramatyce $\mathcal{G}$ to drzewo etykietowane $A \cup N \cup \{\varepsilon\}$ t.ż.

- S jest etykietą korzenia,
- każdy wierzchołek o etykiecie $\in A \cup \{\varepsilon\}$ jest liściem,
- każdy wierzchołek o etykiecie $X \in N$ jest wewnętrzny, oraz



$$X \rightarrow_{\mathcal{G}} X_1 X_2 \dots X_n \quad (n \geq 1)$$

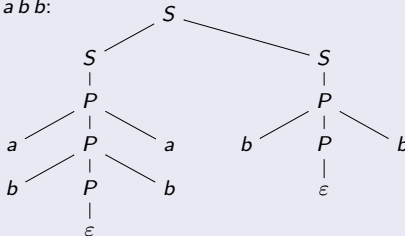
$$X \rightarrow_{\mathcal{G}} \varepsilon$$

Przykład

Drzewo wyprowadzenia dla $a b b a b b$:

$$S \rightarrow S S \mid P$$

$$P \rightarrow a P a \mid b P b \mid \varepsilon$$



Wyprowadzenia a drzewa wyprowadzeń

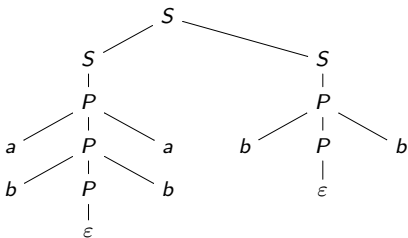
Plon drzewa wyprowadzenia to słowo $w \in A^*$ otrzymane z etykiet liści, w kolejności leksykograficznej (od lewej do prawej), pomijając ε .

Fakt

Niech $w \in A^*$.

$S \rightarrow_G w \iff w$ jest plonem drzewa wyprowadzenia w gramatyce $\mathcal{G}$.

$S \rightarrow_G S S \rightarrow_G S P \rightarrow_G S b P b \rightarrow_G S b b \rightarrow_G$
 $P b b \rightarrow_G a P a b b \rightarrow_G a b P b a b b \rightarrow_G a b b a b b$

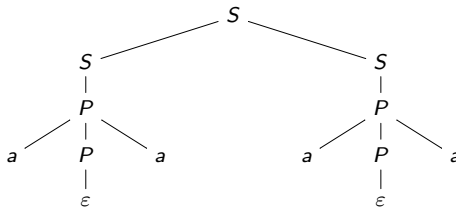
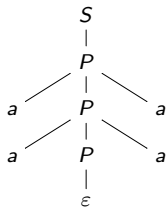


rysunek

Pytanie

Czy każde słowo ma co najwyżej jedno drzewo wyprowadzenia w gramatyce?

Nie, np.



Pytanie

Czy dla każdej gramatyki istnieje równoważna gramatyka *jednoznaczna*, tzn. taka, w której każde słowo ma co najwyżej jedno drzewo wyprowadzenia?

Nie, oto *język niejednoznaczny*:

$$L = \{ a^n b^n c^m d^m : n, m \in \mathbb{N} \} \cup \{ a^n b^m c^m d^n : n, m \in \mathbb{N} \}$$

Przykład

Język palindromów parzystych nad $A = \{a, b\}$, generowany przez gramatykę

$$P \rightarrow_G a P a \mid b P b \mid \varepsilon,$$

jest *najmniejszym* językiem $L \subseteq A^*$ spełniającym warunek:

$$L \supseteq \{a\}L\{a\} \cup \{b\}L\{b\} \cup \{\varepsilon\}.$$

Czyli L jest najmniejszym punktem stałym funkcji monotonicznej

$$L \mapsto \{a\}L\{a\} \cup \{b\}L\{b\} \cup \{\varepsilon\} : \mathcal{P}(A^*) \rightarrow \mathcal{P}(A^*).$$

Pytanie

Jak tę obserwację uogólnić do gramatyk o więcej niż jednym nieterminalu?

- 1 Gramatyki bezkontekstowe - po co?
- 2 Gramatyki bezkontekstowe
- 3 Języki bezkontekstowe a drzewa
- 4 Postaci normalne i problemy decyzyjne

Założmy, że $\varepsilon \notin L(\mathcal{G})$.

$(S \rightarrow \varepsilon)$

Postać normalna Chomsky'ego

$$X \rightarrow a \quad (a \in A) \qquad X \rightarrow YZ \quad (Y, Z \in N)$$



Dowód:

$$X \rightarrow a \quad (a \in A) \qquad X \rightarrow w \quad (w \in N^*)$$

Postać normalna Greibach

$$X \rightarrow a w \quad (a \in A, w \in N^*)$$



Pustość

Dane: gramatyka bezkontekstowa $\mathcal{G}$ (np. w postaci Chomsky'ego)

Wynik: czy $L(\mathcal{G}) = \emptyset$?

Uniwersalność

Dane: gramatyka bezkontekstowa $\mathcal{G}$

Wynik: czy $L(\mathcal{G}) = A^*$?

Równość

Dane: gramatyki bezkontekstowe $\mathcal{G}, \mathcal{G}'$

Wynik: czy $L(\mathcal{G}) = L(\mathcal{G}')$?

Analiza składniowa

Dane: gramatyka bezkontekstowa $\mathcal{G} = (A, N, S, \alpha)$ i słowo $a_1 \dots a_n$
Wynik: czy $a_1 \dots a_n \in L(\mathcal{G})$?

oblicz: $V_{i,j} = \{X \in N : X \twoheadrightarrow_{\mathcal{G}} a_{i+1} \dots a_j\} \quad (0 \leq i < j \leq n)$

rysunek

Algorytm

dla $0 \leq i < j \leq n$,

$V_{i,j} := (i+1 = j) ? \{X \in N : X \rightarrow_{\mathcal{G}} a_j\} : \emptyset$

dla $0 \leq i < j \leq n$,

dla $i < k < j$,

$V_{i,j} := V_{i,j} \cup \{X \in N : \exists X \rightarrow YZ \in \alpha. Y \in V_{i,k}, Z \in V_{k,j}\}$

wynik $:= (S \in V_{0,n})$

W następnym odcinku:

automaty z ...

