

## Zadania – gramatyki bezkontekstowe – jednoznaczność gramatyk

Zad. 1. Rozważamy gramatykę rozszerzonych wyrażeń regularnych:

$$R \rightarrow R' | R | RR | R^* | R^+ | R^? | (R) | a | b | '\epsilon' | '\emptyset'$$

Wykaż, że poniższa gramatyka jest niejednoznaczna – odpowiedź dokładnie uzasadnij. Podaj gramatykę jednoznaczłą generującą ten sam język zachowującą właściwe cechy operatorów.

Zad. 2. Czasami modyfikuje się zapis klasycznej gramatyki bezkontekstowej używając w prawych stronach produkcji gramatyki operatorów podobnych do operatorów dla języków regularnych. Nawiasów kwadratowych można używać do zapisaania opcjonalnej części produkcji. Przykładowo, moglibyśmy napisać:

$$instr \rightarrow \mathbf{if} \text{ wyr } \mathbf{then} \text{ instr } [ \mathbf{else} \text{ instr } ]$$

aby opisać opcjonalną część **else**. Uogólniając,  $A \rightarrow \alpha[\beta]\gamma$  jest równoważne dwóm produkcjom:  $A \rightarrow \alpha\beta\gamma$  oraz  $A \rightarrow \alpha\gamma$ . Nawiasy klamrowe mogą być użyte do opisanie frazy, która powtarza się zero lub więcej razy. Na przykład:

$$instr \rightarrow \mathbf{begin} \text{ instr } \{ ; \text{ instr } \} \mathbf{end}$$

opisuje listę oddzielonych średnikami instrukcji zawartych pomiędzy **begin** i **end**.

Uogólniając,  $A \rightarrow \alpha\{\beta\}\gamma$  jest równoważne  $A \rightarrow \alpha B\gamma$  oraz  $B \rightarrow \beta B | \epsilon$ . W pewnym sensie  $[\beta]$  oznacza wyrażenie regularne  $\beta|\epsilon$ , zaś  $\{\beta\}$  oznacza  $\beta^*$ . Notację tę możemy uogólnić tak, aby po prawej stronie produkcji umieszczać dowolne wyrażenie regularne zbudowane z symboli gramatyki.

- Podaj zbiór produkcji klasycznej gramatyki bezkontekstowej generujący taki sam zbiór napisów, co  $A \rightarrow (b | C) A^* c$ ,
- Podaj zbiór produkcji klasycznej gramatyki bezkontekstowej generujący taki sam zbiór napisów, co  $A \rightarrow b A^* (B | c)$ ,
- Pokaż jak zastąpić dowolną produkcję  $A \rightarrow r$ , gdzie  $r$  jest wyrażeniem regularnym skończonym zbiorem produkcji bezkontekstowych.

Zad. 3. Poniższa gramatyka generuje deklarację pewnego identyfikatora:

$$\begin{aligned} instr &\rightarrow \mathbf{declare} \text{ id } lista\_opcji \\ lista\_opcji &\rightarrow lista\_opcji \text{ opcja } | \epsilon \\ opcja &\rightarrow \text{przecinek} | \text{dokładność} | \text{podstawa} \\ \text{przecinek} &\rightarrow \mathbf{fixed} | \mathbf{floating} \\ \text{dokładność} &\rightarrow \mathbf{single} | \mathbf{double} \\ \text{podstawa} &\rightarrow \mathbf{binary} | \mathbf{decimal} \end{aligned}$$

Powyższa gramatyka może tworzyć nadmiarowe lub sprzeczne deklaracje, takie jak:

$$\mathbf{declare} \text{ id } \mathbf{fixed} \text{ single } \mathbf{fixed} \text{ double}$$

Składnia języka mogłaby zabraniać takich deklaracji. Zmodyfikuj powyższą gramatykę tak, aby każda opcja mogła wystąpić co najwyżej jeden raz, zaś kolejność opcji była dowolna. Ile dodatkowych symboli nieterminalnych musisz wykorzystać?

Zad. 4. Poniższa gramatyka w notacji BNF opisuje fragment składni języka C:

$$\begin{aligned} \langle \text{deklaracja} \rangle &::= \langle \text{typ} \rangle \langle \text{deklarator} \rangle ; \\ \langle \text{typ} \rangle &::= \mathbf{int} | \mathbf{char} \\ \langle \text{deklarator} \rangle &::= * \langle \text{deklarator} \rangle \\ &\quad | \langle \text{deklarator} \rangle [ \text{liczba} ] \\ &\quad | \langle \text{deklarator} \rangle ( \langle \text{typ} \rangle ) \\ &\quad | \underline{\text{nazwa}} \end{aligned}$$

Wykaż, że powyższa gramatyka nie jest jednoznaczna. Przy jakich założeniach lub po jakich zmianach gramatyka byłaby jednoznaczna?