AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

Gramatyki bezkontekstowe, rozbiór gramatyczny

Teoria automatów i języków formalnych

Dr inż. Janusz Majewski
Katedra Informatyki

Gramatyki rekursywne

Niech będzie dana gramatyka bezkontekstowa $G = \langle V, \Sigma, P, S \rangle$.

Gramatyki rekursywne

Gramatykę nazywamy *rekursywną*, jeżeli w gramatyce tej możliwe jest wyprowadzenie $A \Rightarrow^+ \alpha A \beta$ dla pewnego nieterminala $A \in V$, przy czym $\alpha, \beta \in (V \cup \Sigma)^*$.

Gramatykę nazywamy *lewostronnie rekursywną*, jeżeli w gramatyce tej możliwe jest wyprowadzenie $A \Rightarrow^+ A\beta$ dla pewnego nieterminala $A \in V$, przy czym $\beta \in (V \cup \Sigma)^*$.

Gramatykę nazywamy *prawostronnie rekursywną*, jeżeli w gramatyce tej możliwe jest wyprowadzenie $A \Rightarrow^+ \alpha A$ dla pewnego nieterminala $A \in V$, przy czym $\alpha \in (V \cup \Sigma)^*$.

Jeżeli język $L(G)$ jest zbiorem nieskończonym, to jego gramatyka G musi być gramatyką rekursywną.

Frazy

Frazy

Łańcuch δ nazywamy *frazą* formy zdaniowej $\omega = \alpha\delta\beta$ dla symbolu nieterminalnego $A \in V$, wtedy i tylko wtedy, gdy:

$S \Rightarrow^* \alpha A\beta$

$A \Rightarrow^+ \delta$

przy czym $\alpha, \delta, \beta \in (V\cup\Sigma)^*$.

Łańcuch δ nazywamy *frazą prostą* formy zdaniowej $\omega = \alpha\delta\beta$ dla symbolu nieterminalnego $A \in V$, wtedy i tylko wtedy, gdy:

$S \Rightarrow^* \alpha A\beta$

$A \Rightarrow \delta$

przy czym $\alpha, \delta, \beta \in (V\cup\Sigma)^*$.

Osnową formy zdaniowej jest najbardziej na lewo położona fraza prosta (to ostatnie, określenie ma sens w przypadku gramatyk jednoznacznych – patrz dalej).

Wyprowadzenia lewostronne i prawostronne (1)

Wyprowadzenia lewostronne i prawostronne

Forma zdaniowa ψ jest *wyprowadzalna bezpośrednio lewostronnie* z formy zdaniowej ω w gramatyce G, co zapisujemy

$\omega \Rightarrow_{GL} \psi$

jeżeli:

$\omega \Rightarrow_G \psi$

$\omega = \gamma\alpha\delta$

$\psi = \gamma\beta\delta$

$(\alpha \rightarrow \beta) \in P$

$\gamma \in \Sigma^*$

$\alpha, \beta, \delta, \psi, \omega \in (V\cup\Sigma)^*$

Powyższa definicja nie jest ukierunkowana jedynie na gramatyki bezkontekstowe, ale w wyprowadzeniu lewostronnym w gramatyce bezkontekstowej zawsze skrajny lewy nieterminal jest zastępowany prawą stroną pewnej produkcji.

Wyprowadzenia lewostronne i prawostronne (2)

Forma zdaniowa ψ jest *wyprowadzalna bezpośrednio prawostronnie* z formy zdaniowej ω w gramatyce *G*, co zapisujemy

$\omega \Rightarrow_{GP} \psi$

jeżeli:

$\omega \Rightarrow_G \psi$

$\omega = \gamma\alpha\delta$

$\psi = \gamma\beta\delta$

$(\alpha \to \beta) \in P$

$\delta \in \Sigma^*$

$\alpha, \beta, \gamma, \psi, \omega \in (V \cup \Sigma)^*$

Podobnie jak poprzednio, powyższa definicja nie jest ukierunkowana jedynie na gramatyki bezkontekstowe, ale w wyprowadzeniu prawostronnym w gramatyce bezkontekstowej zawsze skrajny prawy nieterminal jest zastępowany prawą stroną pewnej produkcji.

Podobnie jak poprzednio, definiuje się relacje $\Rightarrow_{GL}{}^+$, $\Rightarrow_{GL}{}^*$, $\Rightarrow_{GP}{}^+$, $\Rightarrow_{GP}{}^*$, które są odpowiednio przechodnim oraz przechodnim i zwrotnym domknięciem relacji bezpośredniej wyprowadzalności lewostronnej $\Rightarrow_{GL}$ i prawostronnej $\Rightarrow_{GP}$. Jeżeli wiadomo, o jaką gramatykę chodzi, pomijamy dolny indeks „G" w oznaczeniu tych relacji pisząc po prostu: $\Rightarrow_L{}^+$, $\Rightarrow_L{}^*$, $\Rightarrow_P{}^+$, $\Rightarrow_P{}^*$, $\Rightarrow_L$ oraz $\Rightarrow_P$.

Przykład (1)

<u>Przykład:</u>

Niech będzie dana gramatyka bezkontekstowa $G = \langle V, \Sigma, P, S\rangle$, gdzie:

V = {E, T, F}

Σ = {a, +, *, (,)}

P = {
$E \to E+T \mid T$
$T \to T*F \mid F$
$F \to (E) \mid a$
}

S = E

Formą zdaniową w tej gramatyce jest np. łańcuch:

a+F*T

gdyż:

$E \Rightarrow E+T \Rightarrow T+T \Rightarrow F+T \Rightarrow a+T \Rightarrow a+T*F$

Przykład (2)

$E \Rightarrow E+T \Rightarrow T+T \Rightarrow F+T \Rightarrow a+T \Rightarrow a+T*F$

Powyższe wyprowadzenie polegało na każdorazowym zastępowaniu skrajnego lewego nieterminala prawą stroną jakiejś odpowiedniej produkcji, więc każdy krok tego wyprowadzenia jest wyprowadzeniem lewostronnym. Możemy więc powiedzieć, że rozpatrywany łańcuch jest formą zdaniową wyprowadzalną lewostronnie.

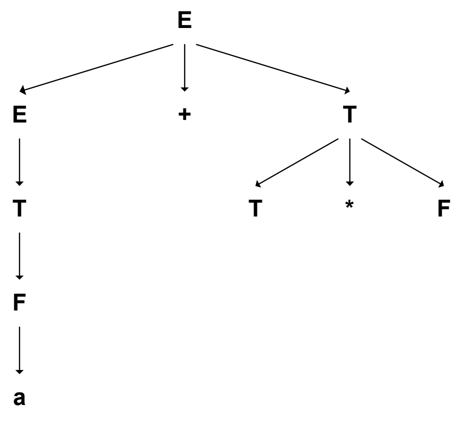
$E \Rightarrow_L E+T \Rightarrow_L T+T \Rightarrow_L F+T \Rightarrow_L a+T \Rightarrow_L a+T*F$

Spróbujmy wyprowadzić badany łańcuch prawostronnie:

$E \Rightarrow_P E+T \Rightarrow_P E+T*F$

Dalsze wyprowadzenie prawostronne wymagałoby zastąpienia nieterminala *F* prawą stroną jakiejś produkcji, ale z uwagi na to, że wyprowadzany łańcuch musi się kończyć właśnie wyprowadzoną sekwencją *+T*F*, nie jest to możliwe, więc *a+F*T* nie jest formą zdaniową wyprowadzalną prawostronnie.

Przykład (3)



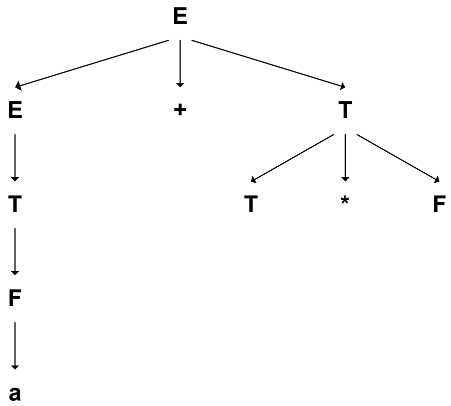
Znajdziemy teraz wszystkie frazy, frazy proste i osnowę analizowanej formy zdaniowej. Rozważymy wyprowadzenia:

$E \Rightarrow^* F+T*F \Rightarrow a+T*F$

$E \Rightarrow^* a+T \Rightarrow a+T*F$

Porównując te wyprowadzenia z odpowiednią definicją widzimy, że *a* jest frazą prostą naszej formy zdaniowej dla nieterminala *F* oraz *T*F* jest frazą prostą naszej formy zdaniowej dla nieterminala *T*. Poza tym *a* jest osnową. Innych fraz prostych rozważana forma zdaniowa nie posiada.

Przykład (4)



Rozważymy teraz wyprowadzenia:

$E \Rightarrow^{*} T+T*F \Rightarrow^{+} a+T*F$

$E \Rightarrow^{*} E+T*F \Rightarrow^{+} a+T*F$

Widać, że *a* jest frazą (ale już nie frazą prostą) dla nieterminali *T* oraz *E*. Badając dalej mamy:

$E \Rightarrow^{*} E \Rightarrow^{+} a+T*F$

Cały łańcuch *a+T*F* jest frazą naszej formy zdaniowej *a+T*F* dla nieterminala *E* stojącego w korzeniu drzewa rozbioru.