Postać normalna Greibach

22 wrześsnia 2006

1 Wprowadzenie

Poniższa tabela przedstawia najczęściej spotykane formy normalne.

Nazwa postaci	Forma produkcji
Postać normalna Greibach (GNF)	$N \to TN^\star$
Słaba postać normalna Chomsky'ego	$N \to T \cup N^\star$
Postać normalna Chomsky'ego (CNF)	$N \to T \cup N^2$
Rozszerzona postać normalna Chomsky'ego	$N \to T \cup N \cup N^2$
Kwadratowa postać normalna Greibach (2-GNF)	$N \to T \cup TN \cup TN^2$
Sześcienna postać normalna Greibach (3-GNF)	$N \to T \cup TN \cup TN^2 \cup TN^3$
Podwójna postać normalna Greibach	$N \to T \cup TN^\star T$
Kwadratowa podwójna postać normalna Greibach	$N \to T \cup TNT \cup TN^2T$
Postać operatorowa	$N \to (T \cup N)^\star - (T \cup N)^\star N^2 (T \cup N)^\star$

1.1 Postać normalna Chomsky'ego

Wszystkie metody z wyjątkiem algorytmu Rosenkrantza zakładają, że gramatyka już jest w postaci normalnej Chomsky'ego lub pewnej odmianie postaci Chomsky'ego.

2 Metoda Paull'a

uporządkuj nieterminale jako $\{ A_1, \ldots, A_n \}$.
for $i := 1$ to n **do**
  for $j := 1$ to $i - 1$ **do**
    for każda produkcja postaci $A_i \to A_j\alpha$ **do**
      usuń produkcję $A_i \to A_j\alpha$ z gramatyki
      for każda produkcja postaci $A_j \to \beta$ **do**
        dodaj $A_i \to \beta\alpha$ do gramatyki
      end for
    end for
  end for
  if A_i jest bezpośrednio lewostronnie rekursywny **then**
    usuń produkcje $A_i \to A_i\alpha_1 \mid \cdots \mid A_i\alpha_r$
    usuń produkcje $A_i \to \beta_1 \mid \cdots \mid \beta_s$ (żadne β_k nie zaczyna się od A_i)
    dodaj produkcje $A_i \to \beta_1 \mid \beta_1 A_i' \mid \cdots \mid \beta_s \mid \beta_s A_i'$
    dodaj produkcje $A_i' \to \alpha_1 \mid \alpha_1 A_i' \mid \cdots \mid \alpha_r \mid \alpha_r A_i'$
  end if
end for
for $i := n$ downto 2 **do**
  for $j := n - 1$ downto 1 **do**
    for każda produkcja postaci $A_j \to A_i\alpha$ **do**
      usuń produkcję $A_j \to A_i\alpha$

for każda produkcja postaci $A_i \to \beta_k$ **do**
dodaj produkcję $A_j \to \beta_k\alpha$
end for
end for
end for
end for
for $i := n$ downto 1 **do**
for $j := n$ downto 1 **do**
for każda produkcja postaci $A'_j \to A_i\alpha$ **do**
usuń produkcję $A'_j \to A_i\alpha$
for każda produkcja postaci $A_i \to \beta_k$ **do**
dodaj produkcję $A'_j \to \beta_k\alpha$
end for
end for
end for
end for

Algorytm generuje gramatyki w postaci normalnej o rozmiarze wykładniczym względem oryginalnej gramatyki.

3 Metoda Rosenkrantz'a

Metoda Rosenkrantz'a wykorzystuje szeregi potęgowe oraz związki między zbiorami produkcji a układami równań.
W metodzie Rosenkrantz'a odchodzi się od tradycyjnych oznaczeń na rzecz notacji używanej w algebrze liniowej. Przykładowo produkcje $X \to aX \mid bY$ będziemy zapisywać $X \to aX + bY$ lub $X = aX + bY$. Wyjątkowo w metodzie nie musi być przestrzegana konwencja w której nieteminale oznaczane są dużymi literami alfabetu łacińskiego. Wygodniejsze często jest oznaczanie nieterminali jak zmiennych, np. $x_1, x_2, \ldots$.
Rozwiązanie równania $y = ay + b$ możemy uzyskać sposób mnemotechniczny. Przekształćmy równanie następująco:
$(1 - a)y = b$
$y = \frac{1}{1-a}b$
$y = (1 + a + a^2 + \ldots)b$
Korzystając z definicji $\star$ Kleene'ego otrzymujemy, że rozwiązaniem równania jest $y = a^*b$.
Z drugiej strony należy pamiętać, że powyższe postępowanie jest mnemotechniczne, uzyskana w podobny sposób odpowiedź $y = ba^*$ nie jest poprawnym rozwiązaniem.
Ponieważ najczęściej z gramatyką związanych jest wiele równań korzysta się z notacji macierzowej.

Sprowadzanie do postaci normalnej Greibach.

Ogólnie gramatykę możemy opisać przy pomocy równania:
$Y = AY + B$, gdzie Y oraz B są wektorami kolumnowymi.
Ponieważ rozważamy tylko gramatyki bezkontekstowe każdy element wektora Y jest pojedynczym nieterminalem.
Zgodnie z powyższym wywodem rozwiązaniem ostatniego równania jest

$Y = A^*B$.
Po dokonaniu transpozycji otrzymujemy równanie
$X = XG + F$, gdzie dla zwięzłości oznaczono $X = Y^T, F = B^T, G = A^T$ oraz jego rozwiązanie $X = FG^*$.
Z własności $\star$ wynika, że $G^\star = I + GG^\star$ a stąd $X = F + FGG^*$.
Oznaczając $H = GG^\star$ możemy zapisać:
$X = FH + F$.
Każdy term wektora F a więc i równania zaczyna się terminalem. Pozostaje tylko znaleźć równania definiujące elementy wektora H.
Na podstawie macierzy G można skonstruować graf zawierający n węzłów i krawędzie z węzła i do węzła j, jeśli $G_{i,j} \neq \varnothing$. Jeśli nie istnieje ścieżka z węzła i do węzła j to w macierzy H należy przyjąć $H_{i,j} = \varnothing$, gdyż $H_{i,j}$ oznacza w grafie zbiór ścieżek z węzła i do węzła j.
Mnożąc lewostronnie związek $G^\star = I + GG^\star$ przez G dostajemy takie równanie:
$H = G + GH$.
Musimy jeszcze sprowadzić ostatnie równanie to postaci normalnej Greibach. W tym celu z macierzy G konstruujemy nową, równoważną macierz K. Jeśli pewien term w macierzy G zaczyna się nieterminalem, nieterminal ten zastępowany jest definicją z równania $X = FH + F$.
W ten sposób otrzymujemy równanie $H = K + KH$ w którym każdym term zaczyna się terminalem.
Mamy więc równania definiujące gramatykę w postaci w której każdy term zaczyna się terminalem:
$X = FH + F$
$H = G + GH$.
Aby otrzymać postać normalną Greibach wystarczy już tylko z każdym terminalem a związać nowy nieterminal C_a oraz produkcję $C_a \to a$ i zastąpić każde wystąpienie terminala a nieterminalem C_a jeśli nie znajduje się on na skrajnie lewej pozycji termu.

Sprowadzanie do podwójnej postaci normalnej Greibach.

Rozwiązaniem równania $x = Ax + b$ jest równanie definiujące $x = A^*b = b + Ab + ABb$, gdzie $B = AA^*$ i może być zdefiniowane równaniem $B = A + AA + ABA$.
Macierz C otrzymywana jest z A poprzez zastąpienie każdego nieterminala kończącego pewien term z A jego definicją z równania definiującego. W ten sposób macierz C równoważna jest macierzy A i każdy term w C zaczyna i kończy się terminalem. Równanie definiujące ma postać: $B = C + AC + ABC$.

Przykład 3.1 Rozważmy gramatykę:

$X_1 \to X_1 a X_2 + b$
$X_2 \to X_2 b X_1 + a$

Zauważmy, że po dokonaniu transpozycji równania mogą być zapisane w postaci

$X = XG + F$

gdzie X i F są wektorami wierszowymi, G jest macierzą kwadratową.
Równania są następujące:

$$\begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} aX_2 & \varnothing \\ \varnothing & bX_1 \end{bmatrix} + \begin{bmatrix} b & a \end{bmatrix}$$

Rozwiązaniem powyższego równania jest $X = FG^\star$. Zauważając , że $G^\star = I + GG^\star$ dostajemy $X = FH + F$ gdzie $H = GG^\star$.

$$\begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} b & a \end{bmatrix} \begin{bmatrix} H_{11} & \varnothing \\ \varnothing & H_{22} \end{bmatrix} + \begin{bmatrix} b & a \end{bmatrix}$$

Obliczamy macierz H, gdzie $H = KH + K$

$$K = \begin{bmatrix} aX_2 & \varnothing \\ \varnothing & bX_1 \end{bmatrix}$$

Mamy:

$H_{11} = aX_2H_{11} + aX_2$
$H_{22} = bX_1H_{22} + bX_1$

Otrzymujemy:

$X_1 = bH_{11} + b$
$X_2 = aH_{22} + a$
$H_{11} = aX_2H_{11} + aX_2$
$H_{22} = bX_1H_{22} + bX_1$

Powyższe równania są w postaci normalnej Greibach. Aby otrzymać podwójną postać normalną Greibach zapiszmy równania w postaci $X = AX + b$. Rozwiązaniem równania jest $X = A^\star b = b + Ab + ABb$, gdzie $B = AA^\star$ i może być zdefiniowane przez $B = C + AC + ABC$, gdzie macierz C otrzymujemy z A przez zastąpienie najbardziej wysuniętych na prawo zmiennych ich odpowiednikami zakończonymi terminalami. Otrzymujemy:

$$\begin{bmatrix} X_1 \\ X_2 \\ H_{11} \\ H_{22} \end{bmatrix} = \begin{bmatrix} \varnothing & \varnothing & b & \varnothing \\ \varnothing & \varnothing & \varnothing & a \\ \varnothing & a & aX_2 & \varnothing \\ b & \varnothing & \varnothing & bX_1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ H_{11} \\ H_{22} \end{bmatrix} + \begin{bmatrix} b \\ a \\ \varnothing \\ \varnothing \end{bmatrix}$$

$$C = \begin{bmatrix} \varnothing & \varnothing & b & \varnothing \\ \varnothing & \varnothing & \varnothing & a \\ \varnothing & a & aa + aaB_{41}b + aaB_{42}a & \varnothing \\ b & \varnothing & \varnothing & bb + bbB_{31}b + bbB_{32}a \end{bmatrix}$$

$$B = \begin{bmatrix} \varnothing & \varnothing & b & \varnothing \\ \varnothing & \varnothing & \varnothing & a \\ \varnothing & a & aa + aaB_{41}b + aaB_{42}a & \varnothing \\ b & \varnothing & \varnothing & bb + bbB_{31}b + bbB_{32}a \end{bmatrix} +$$
$$\begin{bmatrix} \varnothing & \varnothing & b & \varnothing \\ \varnothing & \varnothing & \varnothing & a \\ \varnothing & a & aX_2 & \varnothing \\ b & \varnothing & \varnothing & bX_1 \end{bmatrix} \cdot \begin{bmatrix} \varnothing & \varnothing & b & \varnothing \\ \varnothing & \varnothing & \varnothing & a \\ \varnothing & a & aa + aaB_{41}b + aaB_{42}a & \varnothing \\ b & \varnothing & \varnothing & bb + bbB_{31}b + bbB_{32}a \end{bmatrix} +$$
$$\begin{bmatrix} \varnothing & \varnothing & b & \varnothing \\ \varnothing & \varnothing & \varnothing & a \\ \varnothing & a & aX_2 & \varnothing \\ b & \varnothing & \varnothing & bX_1 \end{bmatrix} \cdot \begin{bmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ B_{41} & B_{42} & B_{43} & B_{44} \end{bmatrix} \cdot \begin{bmatrix} \varnothing & \varnothing & b & \varnothing \\ \varnothing & \varnothing & \varnothing & a \\ \varnothing & a & aa + aaB_{41}b + aaB_{42}a & \varnothing \\ b & \varnothing & \varnothing & bb + bbB_{31}b + bbB_{32}a \end{bmatrix}$$

$$
\begin{bmatrix} B_{11} \\ B_{12} \\ B_{13} \\ B_{14} \\ B_{21} \\ B_{22} \\ B_{23} \\ B_{24} \\ B_{31} \\ B_{32} \\ B_{33} \\ \\ B_{34} \\ \\ B_{41} \\ B_{42} \\ B_{43} \\ \\ B_{44} \\ \\ \end{bmatrix} = \begin{bmatrix} bB_{34}b \\ ba + bB_{33}a \\ b + baa + baaB_{41}b + baaB_{42}a + bB_{31}b + bB_{33}aa + bB_{33}aaB_{41}b + bB_{33}aaB_{42}a \\ bB_{32}a + bB_{34}bb + bB_{34}bbB_{31}b + bB_{34}bbB_{32}a \\ ab + aB_{44}b \\ aB_{43}a \\ aB_{41}b + aB_{43}aa + aB_{43}aaB_{41}b + aB_{43}aaB_{42}a \\ a + abb + abbB_{31}b + abbB_{32}a + aB_{42}a + aB_{44}bb + aB_{44}bbB_{31}b + aB_{44}bbB_{32}a \\ aB_{24}b + aX_2B_{34}b \\ a + aX_2a + aB_{23}a + aX_2B_{33}a \\ aa + aaB_{41}b + aaB_{42}a + aX_2aa + aX_2aaB_{41}b + aX_2aaB_{42}a + \\ aB_{21}b + aX_2B_{31}b + aB_{23}aa + aB_{23}aaB_{41}b + aB_{23}aaB_{42}a + aX_2B_{33}aa + aX_2B_{33}aaB_{41}b + aX_2B_{33}aaB_{42}a \\ aa + \\ aB_{22}a + aX_2B_{32}a + aB_{24}bb + aB_{24}bbB_{31}b + aB_{24}bbB_{32}a + aX_2B_{34}bb + aX_2B_{34}bbB_{31}b + aX_2B_{34}bbB_{32}a \\ b + bX_1b + bB_{14}b + bX_1B_{44}b \\ bB_{13}a + bX_1B_{43}a \\ bb + \\ bB_{11}b + bX_1B_{41}b + bB_{13}aa + bB_{13}aaB_{41}b + bB_{13}aaB_{42}a + bX_1B_{43}aa + bX_1B_{43}aaB_{41}b + bX_1B_{43}aaB_{42}a \\ bb + bbB_{31}b + bbB_{32}a + bX_1bb + bX_1bbB_{31}b + bX_1bbB_{32}a + \\ bB_{12}a + bX_1B_{42}a + bB_{14}bb + bB_{14}bbB_{31}b + bB_{14}bbB_{32}a + bX_1B_{44}bb + bX_1B_{44}bbB_{31}b + bX_1B_{44}bbB_{32}a \end{bmatrix}
$$

Ostatecznie, uwzględniając, że symbole H_{11}, H_{22} są nieosiągalne, otrzymujemy gramatykę:

$X_1 \to b + bB_{31}b + bB_{32}a$
$X_2 \to a + aB_{41}b + aB_{42}a$
$B_{11} \to bB_{34}b$
$B_{12} \to ba + bB_{33}a$
$B_{13} \to b + baa + baaB_{41}b + baaB_{42}a + bB_{31}b + bB_{33}aa + bB_{33}aaB_{41}b + bB_{33}aaB_{42}a$
$B_{14} \to bB_{32}a + bB_{34}bb + bB_{34}bbB_{31}b + bB_{34}bbB_{32}a$
$B_{21} \to ab + aB_{44}b$
$B_{22} \to aB_{43}a$
$B_{23} \to aB_{41}b + aB_{43}aa + aB_{43}aaB_{41}b + aB_{43}aaB_{42}a$
$B_{24} \to a + abb + abbB_{31}b + abbB_{32}a + aB_{42}a + aB_{44}bb + aB_{44}bbB_{31}b + aB_{44}bbB_{32}a$
$B_{31} \to aB_{24}b + aX_2B_{34}b$
$B_{32} \to a + aX_2a + aB_{23}a + aX_2B_{33}a$
$B_{33} \to aa + aaB_{41}b + aaB_{42}a + aX_2aa + aX_2aaB_{41}b + aX_2aaB_{42}a +$
$aB_{21}b+aX_2B_{31}b+aB_{23}aa+aB_{23}aaB_{41}b+aB_{23}aaB_{42}a+aX_2B_{33}aa+aX_2B_{33}aaB_{41}b+aX_2B_{33}aaB_{42}a$
$B_{34} \to aa +$
$aB_{22}a+aX_2B_{32}a+aB_{24}bb+aB_{24}bbB_{31}b+aB_{24}bbB_{32}a+aX_2B_{34}bb+aX_2B_{34}bbB_{31}b+aX_2B_{34}bbB_{32}a$
$B_{41} \to b + bX_1b + bB_{14}b + bX_1B_{44}b$
$B_{42} \to bB_{13}a + bX_1B_{43}a$
$B_{43} \to bb +$
$bB_{11}b+bX_1B_{41}b+bB_{13}aa+bB_{13}aaB_{41}b+bB_{13}aaB_{42}a+bX_1B_{43}aa+bX_1B_{43}aaB_{41}b+bX_1B_{43}aaB_{42}a$
$B_{44} \to bb + bbB_{31}b + bbB_{32}a + bX_1bb + bX_1bbB_{31}b + bX_1bbB_{32}a +$
$bB_{12}a+bX_1B_{42}a+bB_{14}bb+bB_{14}bbB_{31}b+bB_{14}bbB_{32}a+bX_1B_{44}bb+bX_1B_{44}bbB_{31}b+bX_1B_{44}bbB_{32}a$

Przykład 3.2 Rozważmy gramatykę:

$S \to SA + b$
$A \to BA + a$
$B \to SA + AB$

Mamy następujące równania:

$$\begin{bmatrix} S & A & B \end{bmatrix} = \begin{bmatrix} S & A & B \end{bmatrix} \begin{bmatrix} A & \varnothing & A \\ \varnothing & \varnothing & B \\ \varnothing & A & \varnothing \end{bmatrix} + \begin{bmatrix} b & a & \varnothing \end{bmatrix}$$
$$\begin{bmatrix} S & A & B \end{bmatrix} = \begin{bmatrix} b & a & \varnothing \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ \varnothing & H_{22} & H_{23} \\ \varnothing & H_{32} & H_{33} \end{bmatrix} + \begin{bmatrix} b & a & \varnothing \end{bmatrix}$$

Obliczamy macierz H, gdzie $H = KH + K$

$$K = \begin{bmatrix} bH_{12} + aH_{22} + a & \varnothing & bH_{12} + aH_{22} + a \\ \varnothing & \varnothing & bH_{13} + aH_{23} \\ \varnothing & bH_{12} + aH_{22} + a & \varnothing \end{bmatrix}$$

Mamy:

$H_{11} = bH_{12}H_{11} + aH_{22}H_{11} + aH_{11} + bH_{12} + aH_{22} + a$
$H_{12} = bH_{12}H_{12} + aH_{22}H_{12} + aH_{12} + bH_{12}H_{32} + aH_{22}H_{32} + aH_{32}$
$H_{22} = bH_{13}H_{32} + aH_{23}H_{32}$
$H_{32} = bH_{12}H_{22} + aH_{22}H_{22} + aH_{22} + bH_{12} + aH_{22} + a$

$H_{13} = bH_{12}H_{13} + aH_{22}H_{13} + aH_{13} + bH_{12}H_{33} + aH_{22}H_{33} + aH_{33} + bH_{12} + aH_{22} + a$
$H_{23} = bH_{13}H_{33} + aH_{23}H_{33} + bH_{13} + aH_{23}$
$H_{33} = bH_{12}H_{23} + aH_{22}H_{23} + aH_{23}$

Otrzymujemy:

$S = bH_{11} + b$
$A = bH_{12} + aH_{22} + a$
$B = bH_{13} + aH_{23}$
$H_{11} = bH_{12}H_{11} + aH_{22}H_{11} + aH_{11} + bH_{12} + aH_{22} + a$
$H_{12} = bH_{12}H_{12} + aH_{22}H_{12} + aH_{12} + bH_{12}H_{32} + aH_{22}H_{32} + aH_{32}$
$H_{22} = bH_{13}H_{32} + aH_{23}H_{32}$
$H_{32} = bH_{12}H_{22} + aH_{22}H_{22} + aH_{22} + bH_{12} + aH_{22} + a$
$H_{13} = bH_{12}H_{13} + aH_{22}H_{13} + aH_{13} + bH_{12}H_{33} + aH_{22}H_{33} + aH_{33} + bH_{12} + aH_{22} + a$
$H_{23} = bH_{13}H_{33} + aH_{23}H_{33} + bH_{13} + aH_{23}$
$H_{33} = bH_{12}H_{23} + aH_{22}H_{23} + aH_{23}$

Po usunięciu symboli nieosiągalnych otrzymujemy:

$S = bH_{11} + b$
$H_{11} = bH_{12}H_{11} + aH_{22}H_{11} + aH_{11} + bH_{12} + aH_{22} + a$
$H_{12} = bH_{12}H_{12} + aH_{22}H_{12} + aH_{12} + bH_{12}H_{32} + aH_{22}H_{32} + aH_{32}$
$H_{22} = bH_{13}H_{32} + aH_{23}H_{32}$
$H_{32} = bH_{12}H_{22} + aH_{22}H_{22} + aH_{22} + bH_{12} + aH_{22} + a$
$H_{13} = bH_{12}H_{13} + aH_{22}H_{13} + aH_{13} + bH_{12}H_{33} + aH_{22}H_{33} + aH_{33} + bH_{12} + aH_{22} + a$
$H_{23} = bH_{13}H_{33} + aH_{23}H_{33} + bH_{13} + aH_{23}$
$H_{33} = bH_{12}H_{23} + aH_{22}H_{23} + aH_{23}$

Przykład 3.3 Rozważmy gramatykę:

$S \to AB + BS$
$A \to BA + a$
$B \to AS + b$

Mamy następujące równania:

$$\begin{bmatrix} S & A & B \end{bmatrix} = \begin{bmatrix} S & A & B \end{bmatrix} \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ B & \varnothing & S \\ S & A & \varnothing \end{bmatrix} + \begin{bmatrix} \varnothing & a & b \end{bmatrix}$$

$$\begin{bmatrix} S & A & B \end{bmatrix} = \begin{bmatrix} \varnothing & a & b \end{bmatrix} \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} + \begin{bmatrix} \varnothing & a & b \end{bmatrix}$$

Obliczamy macierz H, gdzie $H = KH + K$

$$K = \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ aH_{23} + bH_{33} + b & \varnothing & aH_{21} + bH_{31} \\ aH_{21} + bH_{31} & aH_{22} + bH_{32} + a & \varnothing \end{bmatrix}$$

Mamy:

$H_{21} = aH_{21}H_{31} + bH_{31}H_{31} + aH_{23} + bH_{33} + b$
$H_{31} = aH_{22}H_{21} + bH_{32}H_{21} + aH_{21} + bH_{31}$
$H_{22} = aH_{21}H_{32} + bH_{31}H_{32}$
$H_{32} = aH_{22}H_{22} + bH_{32}H_{22} + aH_{22} + bH_{32} + a$
$H_{23} = aH_{21}H_{33} + bH_{31}H_{33} + aH_{21} + bH_{31}$
$H_{33} = aH_{22}H_{23} + bH_{32}H_{23} + aH_{23}$

Otrzymujemy:

$S = aH_{21} + bH_{31}$
$A = aH_{22} + bH_{32} + a$
$B = aH_{23} + bH_{33} + b$
$H_{21} = aH_{21}H_{31} + bH_{31}H_{31} + aH_{23} + bH_{33} + b$
$H_{31} = aH_{22}H_{21} + bH_{32}H_{21} + aH_{21} + bH_{31}$
$H_{22} = aH_{21}H_{32} + bH_{31}H_{32}$
$H_{32} = aH_{22}H_{22} + bH_{32}H_{22} + aH_{22} + bH_{32} + a$
$H_{23} = aH_{21}H_{33} + bH_{31}H_{33} + aH_{21} + bH_{31}$
$H_{33} = aH_{22}H_{23} + bH_{32}H_{23} + aH_{23}$

Po usunięciu symboli nieosiągalnych otrzymujemy:

$S = aH_{21} + bH_{31}$
$H_{21} = aH_{21}H_{31} + bH_{31}H_{31} + aH_{23} + bH_{33} + b$
$H_{31} = aH_{22}H_{21} + bH_{32}H_{21} + aH_{21} + bH_{31}$
$H_{22} = aH_{21}H_{32} + bH_{31}H_{32}$
$H_{32} = aH_{22}H_{22} + bH_{32}H_{22} + aH_{22} + bH_{32} + a$
$H_{23} = aH_{21}H_{33} + bH_{31}H_{33} + aH_{21} + bH_{31}$
$H_{33} = aH_{22}H_{23} + bH_{32}H_{23} + aH_{23}$

4 Metoda Autebert'a-Berstel'a-Boasson'a

Metoda ta jest także stosunkowo skomplikowana. W metodzie zakładamy, że gramatyka jest już w słabej postaci normalnej Chomsky'ego. Dla każdego nieterminala X zdefiniujmy:

$L(a, X) = a^{-1}L(X) = \{\, m \in N^\star \mid X \xrightarrow[lm]{\star} \alpha \to am,\ \alpha \in N^\star \,\}$
$R(X, a) = R(X)a^{-1} = \{\, m \in N^\star \mid X \xrightarrow[rm]{\star} \alpha \to ma,\ \alpha \in N^\star \,\}$
Oznaczmy też:

$\mathcal{L} = \{\, L(a, X) \mid a \in T,\ X \in N \,\}$
$\mathcal{R} = \{\, R(a, X) \mid a \in T,\ X \in N \,\}$

Zbiory $L(a, X)$, $R(X, a)$ są nieskończone, jednakże zbiory $\mathcal{L}$, $\mathcal{R}$ są już skończone. W związku z tym także $\mathcal{H}$ t.j. domknięcie $\mathcal{L} \bigcup \mathcal{R}$ względem lewostronnego oraz prawostronnego ilorazu przez symbole terminalne jest zbiorem skończonym.

Konstrukcja gramatyki w postaci normalnej Greibach.

1. Każda produkcja terminalna oryginalnej gramatyki jest produkcją nowej gramatyki w postaci normalnej Greibach.
2. Dla każdej zmiennej $X \in N$ oryginalnej gramatyki dodaj (skończony) zbiór produkcji $X \to aL$, gdzie $a \in T,\ L = L(a, X) \in \mathcal{H}$
3. Da każdej nowej zmiennej $L \in \mathcal{H}$ dodaj (skończony) zbiór produkcji $L \to XL'$, gdzie $X \in N,\ X^{-1}L \in \mathcal{H}$. Jeśli $\epsilon \in L$ dodaj produkcję $L \to \epsilon$.

4. W każdej nowej produkcji dodanej powyżej nieterminal rozpoczynający prawą stronę produkcji zastąp prawymi stronami wygenerowanymi w punkcie 2.

Konstrukcja gramatyki w sześciennej podwójnej postaci normalnej Greibach.

1. Każdy język $L \in \mathcal{H}$ zapisz jako sumę składników $L'X$, dla każdego $X \in V$, gdzie $L' = LX^{-1}$. Dodaj do sumy X jeśli $\epsilon \in L'$.
   Zastąp X sumą składników Ra, gdzie R jest zmienną związaną z językiem $R(X,a)$. W sumie uzwględnij a jeśli $\epsilon \in R(X,a)$.
2. Zastąp w gramatyce w postaci normalnej Greibach każdy skrajnie prawy nieterminal wyrażeniem otrzymanym w punkcie 1.

Konstrukcja gramatyki w kwadratowej podwójnej postaci normalnej Greibach.

Do rodziny zbiorów $\mathcal{H}$ dodaj zmienne $Y \in W$ reprezentujące $L \cdot L' \in \mathcal{H}\mathcal{H}$, gdzie L opisywany jest lewostronnym ilorazem, zaś L' prawostronnym ilorazem. Każdy iloczyn 4 kolejnych nieterminali zastąp iloczynem 2 elementów z W. Każdy iloczyn 3 kolejnych nieterminali zastąp iloczynem elementu z W i elementu z $\mathcal{H}$.

Przykład 4.1 Ponownie rozważmy gramatykę:

$X_1 \to X_1 a X_2 + b$
$X_2 \to X_2 b X_1 + a$

Na początku przekształcamy gramatykę do słabej postaci normalnej Chomsky'ego.

$X_1 \to X_1 A X_2 + b$
$X_2 \to X_2 B X_1 + a$
$A \to a$
$B \to b$

Rozpatrzmy rodziny $\mathcal{L}$ i $\mathcal{R}$.

$$
\begin{array}{rlrl}
L(a,X_1) = & \varnothing & R(X_1,a) = & (X_1AX_2B)^\star X_1A = L_2 \\
L(b,X_1) = & (AX_2)^\star = L_0 & R(X_1,b) = & (X_1AX_2B)^\star = L_3 \\
L(a,X_2) = & (BX_1)^\star = L_1 & R(X_2,a) = & (X_2BX_1A)^\star = L_4 \\
L(b,X_2) = & \varnothing & R(X_2,b) = & (X_2BX_1A)^\star X_2B = L_5 \\
L(a,A) = & \{\epsilon\} = E & R(A,a) = & \{\epsilon\} = E \\
L(b,A) = & \varnothing & R(A,b) = & \varnothing \\
L(a,B) = & \varnothing & R(B,a) = & \varnothing \\
L(b,B) = & \{\epsilon\} = E & R(B,b) = & \{\epsilon\} = E
\end{array}
$$

Obliczmy domknięcie $\mathcal{L} \bigcup \mathcal{R}$ względem lewostronnego oraz prawostronnego ilorazu.

$$
\begin{array}{rlrlrl}
A^{-1}L_0 = & X_2(AX_2)^\star & = L_6 & L_0X_2^{-1} = & (AX_2)^\star A & = L_7 \\
B^{-1}L_1 = & X_1(BX_1)^\star & = L_8 & L_1X_1^{-1} = & (BX_1)^\star B & = L_9 \\
X_1^{-1}L_2 = & A(X_2BX_1A)^\star & = L_{10} & L_2A^{-1} = & (X_1AX_2B)^\star X_1 & = L_{11} \\
X_1^{-1}L_3 = & AX_2B(X_1AX_2B)^\star & = L_{12} & L_3B^{-1} = & (X_1AX_2B)^\star X_1AX_2 & = L_{13} \\
X_2^{-1}L_4 = & BX_1A(X_2BX_1A)^\star & = L_{14} & L_4A^{-1} = & (X_2BX_1A)^\star X_2BX_1 & = L_{15} \\
X_2^{-1}L_5 = & B(X_1AX_2B)^\star & = L_{16} & L_5B^{-1} = & (X_2BX_1A)^\star X_2 & = L_{17} \\
X_2^{-1}L_6 = & (AX_2)^\star & = L_0 & L_6X_2^{-1} = & (X_2A)^\star & = L_{18} \\
A^{-1}L_7 = & (X_2A)^\star & = L_{18} & L_7A^{-1} = & (AX_2)^\star & = L_0 \\
X_1^{-1}L_8 = & (BX_1)^\star & = L_1 & L_8X_1^{-1} = & (X_1B)^\star & = L_{19} \\
B^{-1}L_9 = & (X_1B)^\star & = L_{19} & L_9B^{-1} = & (BX_1)^\star & = L_1 \\
A^{-1}L_{10} = & (X_2BX_1A)^\star & = L_4 & L_{10}A^{-1} = & (AX_2BX_1)^\star & = L_{20} \\
X_1^{-1}L_{11} = & (AX_2BX_1)^\star & = L_{20} & L_{11}X_1^{-1} = & (X_1AX_2B)^\star & = L_3 \\
A^{-1}L_{12} = & X_2B(X_1AX_2B)^\star & = L_5 & L_{12}B^{-1} = & (BX_1AX_2)^\star & = L_{21} \\
X_1^{-1}L_{13} = & (AX_2BX_1)^\star AX_2 & = L_{21} & L_{13}X_2^{-1} = & (X_1AX_2B)^\star X_1A & = L_2 \\
B^{-1}L_{14} = & X_1A(X_2BX_1A)^\star & = L_2 & L_{14}A^{-1} = & (AX_2BX_1)^\star & = L_{22} \\
X_2^{-1}L_{15} = & (BX_1AX_2)^\star BX_1 & = L_{22} & L_{15}X_1^{-1} = & (X_2BX_1A)^\star X_2B & = L_5 \\
B^{-1}L_{16} = & (X_1AX_2B)^\star & = L_3 & L_{16}B^{-1} = & (BX_1AX_2)^\star & = L_{23} \\
X_2^{-1}L_{17} = & (BX_1AX_2)^\star & = L_{23} & L_{17}X_2^{-1} = & (X_2BX_1A)^\star & = L_4 \\
X_2^{-1}L_{18} = & A(X_2A)^\star & = L_7 & L_{18}A^{-1} = & (X_2A)^\star X_2 & = L_6 \\
X_1^{-1}L_{19} = & B(X_1B)^\star & = L_9 & L_{19}B^{-1} = & (X_1B)^\star X_1 & = L_8 \\
A^{-1}L_{20} = & X_2BX_1(AX_2BX_1)^\star & = L_{15} & L_{20}X_1^{-1} = & (AX_2BX_1)^\star AX_2B & = L_{12} \\
A^{-1}L_{21} = & X_2(BX_1AX_2)^\star & = L_{17} & L_{21}X_2^{-1} = & (AX_2BX_1)^\star A & = L_{10} \\
B^{-1}L_{22} = & X_1(AX_2BX_1)^\star & = L_{11} & L_{22}X_1^{-1} = & (BX_1AX_2)^\star B & = L_{16} \\
B^{-1}L_{23} = & X_1AX_2(BX_1AX_2)^\star & = L_{13} & L_{23}X_2^{-1} = & (BX_1AX_2)^\star BX_1A & = L_{14}
\end{array}
$$

Równoważna gramatyka wygląda następująco:

$$
\begin{array}{rlrl}
X_1 &\to bL_0 & X_1 &\to L_2a + L_3b \\
X_2 &\to aL_1 & X_2 &\to L_4a + L_5b \\
A &\to aE & A &\to Ea \\
B &\to bE & B &\to Eb \\
E &\to \epsilon & E &\to \epsilon \\
L_0 &\to AL_6 + \epsilon & L_0 &\to L_7X_2 + \epsilon \\
L_1 &\to BL_8 + \epsilon & L_1 &\to L_9X_1 + \epsilon \\
L_2 &\to X_1L_{10} & L_2 &\to L_{11}A \\
L_3 &\to X_1L_{12} + \epsilon & L_3 &\to L_{13}B + \epsilon \\
L_4 &\to X_2L_{14} + \epsilon & L_4 &\to L_{15}A + \epsilon \\
L_5 &\to X_2L_{16} & L_5 &\to L_{17}B \\
L_6 &\to X_2L_0 & L_6 &\to L_{18}X_2 \\
L_7 &\to AL_{18} & L_7 &\to L_0A \\
L_8 &\to X_1L_1 & L_8 &\to L_{19}X_1 \\
L_9 &\to BL_{19} & L_9 &\to L_1B \\
L_{10} &\to AL_4 & L_{10} &\to L_{20}A \\
L_{11} &\to X_1L_{20} & L_{11} &\to L_3X_1 \\
L_{12} &\to AL_5 & L_{12} &\to L_{21}B \\
L_{13} &\to X_1L_{21} & L_{13} &\to L_2X_2 \\
L_{14} &\to BL_2 & L_{14} &\to L_{22}A \\
L_{15} &\to X_2L_{22} & L_{15} &\to L_5X_1 \\
L_{16} &\to BL_3 & L_{16} &\to L_{23}B \\
L_{17} &\to X_2L_{23} & L_{17} &\to L_4X_2 \\
L_{18} &\to X_2L_7 + \epsilon & L_{18} &\to L_6A + \epsilon \\
L_{19} &\to X_1L_9 + \epsilon & L_{19} &\to L_8B + \epsilon \\
L_{20} &\to AL_{15} + \epsilon & L_{20} &\to L_{12}X_1 + \epsilon \\
L_{21} &\to AL_{17} & L_{21} &\to L_{10}X_2 \\
L_{22} &\to BL_{11} & L_{22} &\to L_{16}X_1 \\
L_{23} &\to BL_{13} + \epsilon & L_{23} &\to L_{14}X_2 + \epsilon
\end{array}
$$

Zastępując X_1, X_2, A i B otrzymujemy:

$$
\begin{array}{rl}
X_1 \rightarrow bL_0 & X_1 \rightarrow L_2a + L_3b \\
X_2 \rightarrow aL_1 & X_2 \rightarrow L_4a + L_5b \\
A \rightarrow aE & A \rightarrow Ea \\
B \rightarrow bE & B \rightarrow Eb \\
E \rightarrow \epsilon & E \rightarrow \epsilon \\
L_0 \rightarrow aEL_6 + \epsilon & L_0 \rightarrow L_7L_4a + L_7L_5b + \epsilon \\
L_1 \rightarrow bEL_8 + \epsilon & L_1 \rightarrow L_9L_2a + L_9L_3b + \epsilon \\
L_2 \rightarrow bL_0L_{10} & L_2 \rightarrow L_{11}Ea \\
L_3 \rightarrow bL_0L_{12} + \epsilon & L_3 \rightarrow L_{13}Eb + \epsilon \\
L_4 \rightarrow aL_1L_{14} + \epsilon & L_4 \rightarrow L_{15}Ea + \epsilon \\
L_5 \rightarrow aL_1L_{16} & L_5 \rightarrow L_{17}Eb \\
L_6 \rightarrow aL_1L_0 & L_6 \rightarrow L_{18}L_4a + L_{18}L_5b \\
L_7 \rightarrow aEL_{18} & L_7 \rightarrow L_0Ea \\
L_8 \rightarrow bL_0L_1 & L_8 \rightarrow L_{19}L_2a + L_{19}L_3b \\
L_9 \rightarrow bEL_{19} & L_9 \rightarrow L_1Eb \\
L_{10} \rightarrow aEL_4 & L_{10} \rightarrow L_{20}Ea \\
L_{11} \rightarrow bL_0L_{20} & L_{11} \rightarrow L_3L_2a + L_3L_3b \\
L_{12} \rightarrow aEL_5 & L_{12} \rightarrow L_{21}Eb \\
L_{13} \rightarrow bL_0L_{21} & L_{13} \rightarrow L_2L_4a + L_2L_5b \\
L_{14} \rightarrow bEL_2 & L_{14} \rightarrow L_{22}Ea \\
L_{15} \rightarrow aL_1L_{22} & L_{15} \rightarrow L_5L_2a + L_5L_3b \\
L_{16} \rightarrow bEL_3 & L_{16} \rightarrow L_{23}Eb \\
L_{17} \rightarrow aL_1L_{23} & L_{17} \rightarrow L_4L_4a + L_4L_5b \\
L_{18} \rightarrow aL_1L_7 + \epsilon & L_{18} \rightarrow L_6Ea + \epsilon \\
L_{19} \rightarrow bL_0L_9 + \epsilon & L_{19} \rightarrow L_8Eb + \epsilon \\
L_{20} \rightarrow aEL_{15} + \epsilon & L_{20} \rightarrow L_{12}L_2a + L_{12}L_3b + \epsilon \\
L_{21} \rightarrow aEL_{17} & L_{21} \rightarrow L_{10}L_4a + L_{10}L_5b \\
L_{22} \rightarrow bEL_{11} & L_{22} \rightarrow L_{16}L_2a + L_{16}L_3b \\
L_{23} \rightarrow bEL_{13} + \epsilon & L_{23} \rightarrow L_{14}L_4a + L_{14}L_5b + \epsilon
\end{array}
$$

Po usunięciu ϵ-produkcji mamy:

$$
\begin{aligned}
X_1 &\to bL_0 + b \\
X_2 &\to aL_1 + a \\
A &\to a \\
B &\to b \\
L_0 &\to aL_6 \\
L_1 &\to bL_8 \\
L_2 &\to bL_0L_{10} + bL_{10} \\
L_3 &\to bL_0L_{12} + bL_{12} \\
L_4 &\to aL_1L_{14} + aL_{14} \\
L_5 &\to aL_1L_{16} + aL_{16} \\
L_6 &\to aL_1L_0 + aL_1 + aL_0 + a \\
L_7 &\to aL_{18} + a \\
L_8 &\to bL_0L_1 + bL_0 + bL_1 + b \\
L_9 &\to bL_{19} + b \\
L_{10} &\to aL_4 + a \\
L_{11} &\to bL_0L_{20} + bL_0 + bL_{20} + b \\
L_{12} &\to aL_5 \\
L_{13} &\to bL_0L_{21} + bL_{21} \\
L_{14} &\to bL_2 \\
L_{15} &\to aL_1L_{22} + aL_{22} \\
L_{16} &\to bL_3 + b \\
L_{17} &\to aL_1L_{23} + aL_1 + aL_{23} + a \\
L_{18} &\to aL_1L_7 + aL_7 \\
L_{19} &\to bL_0L_9 + bL_9 \\
L_{20} &\to aL_{15} \\
L_{21} &\to aL_{17} \\
L_{22} &\to bL_{11} \\
L_{23} &\to bL_{13}
\end{aligned}
\qquad
\begin{aligned}
X_1 &\to L_2a + L_3b + b \\
X_2 &\to L_4a + L_5b + a \\
A &\to a \\
B &\to b \\
L_0 &\to L_7L_4a + L_7a + L_7L_5b \\
L_1 &\to L_9L_2a + L_9L_3b + L_9b \\
L_2 &\to L_{11}a \\
L_3 &\to L_{13}b \\
L_4 &\to L_{15}a \\
L_5 &\to L_{17}b \\
L_6 &\to L_{18}L_4a + L_{18}a + L_4a + a + L_{18}L_5b + L_5b \\
L_7 &\to L_0a + a \\
L_8 &\to L_{19}L_2a + L_2a + L_{19}L_3b + L_{19}b + L_3b + b \\
L_9 &\to L_1b + b \\
L_{10} &\to L_{20}a + a \\
L_{11} &\to L_3L_2a + L_2a + L_3L_3b + L_3b + b \\
L_{12} &\to L_{21}b \\
L_{13} &\to L_2L_4a + L_2a + L_2L_5b \\
L_{14} &\to L_{22}a \\
L_{15} &\to L_5L_2a + L_5L_3b + L_5b \\
L_{16} &\to L_{23}b + b \\
L_{17} &\to L_4L_4a + L_4a + a + L_4L_5b + L_5b \\
L_{18} &\to L_6a \\
L_{19} &\to L_8b \\
L_{20} &\to L_{12}L_2a + L_{12}L_3b + L_{12}b \\
L_{21} &\to L_{10}L_4a + L_{10}a + L_{10}L_5b \\
L_{22} &\to L_{16}L_2a + L_{16}L_3b + L_{16}b \\
L_{23} &\to L_{14}L_4a + L_{14}a + L_{14}L_5b
\end{aligned}
$$

Zastępujemy najbardziej wysunięte na prawo zmienne w lewej gramatyce produkcjami prawej gramatyki, i usuwamy bezużyteczne produkcje dla A i B:

$X_1 \to bL_7L_4a + bL_7a + bL_7L_5b + b$
$X_2 \to aL_9L_2a + aL_9L_3b + aL_9b + a$
$L_0 \to aL_{18}L_4a + aL_{18}a + aL_4a + a + aL_{18}L_5b + aL_5b$
$L_1 \to bL_{19}L_2a + bL_2a + bL_{19}L_3b + bL_{19}b + bL_3b + bb$
$L_2 \to bL_0L_{20}a + bL_0a + bL_{20}a + ba$
$L_3 \to bL_0L_{21}b + bL_{21}b$
$L_4 \to aL_1L_{22}a + aL_{22}a$
$L_5 \to aL_1L_{23}b + aL_1b + aL_{23}b + ab$
$L_6 \to aL_1L_7L_4a + aL_1L_7a + aL_1L_7L_5b + aL_9L_2a + aL_9L_3b + aL_9b + aL_7L_4a + aL_7a + aL_7L_5b + a$
$L_7 \to aL_6a + a$
$L_8 \to bL_0L_9L_2a + bL_0L_9L_3b + bL_0L_9b + bL_7L_4a + bL_7a + bL_7L_5b + bL_9L_2a + bL_9L_3b + bL_9b + b$
$L_9 \to bL_8b + b$
$L_{10} \to aL_{15}a + a$
$L_{11} \to bL_0L_{12}L_2a + bL_0L_{12}L_3b + bL_0L_3b + bL_7L_4a + bL_7a + bL_7L_5b + bL_{12}L_2a + bL_{12}L_3b + bL_3b + b$
$L_{12} \to aL_{17}b$
$L_{13} \to bL_0L_{10}L_4a + bL_0L_{10}a + bL_0L_{10}L_5b + bL_{10}L_4a + bL_{10}a + bL_{10}L_5b$
$L_{14} \to bL_{11}a$
$L_{15} \to aL_1L_{16}L_2a + aL_1L_{16}L_3b + aL_1L_{16}b + aL_{16}L_2a + aL_{16}L_3b + aL_{16}b$
$L_{16} \to bL_{13}b + b$
$L_{17} \to aL_1L_{14}L_4a + aL_1L_{14}a + aL_1L_{14}L_5b + aL_9L_2a + aL_9L_3b + aL_9b + aL_{14}L_4a + aL_{14}a + aL_{14}L_5b + a$
$L_{18} \to aL_1L_0a + aL_1a + aL_0a + aa$
$L_{19} \to bL_0L_1b + bL_0b + bL_1b + bb$
$L_{20} \to aL_5L_2a + aL_5L_3b + aL_5b$
$L_{21} \to aL_4L_4a + aL_4a + aa + aL_4L_5b + aL_5b$
$L_{22} \to bL_3L_2a + bL_2a + bL_3L_3b + bL_3b + bb$
$L_{23} \to bL_2L_4a + bL_2a + bL_2L_5b$

Przykład 4.2 - poprawka artykułu [2]
Rozważamy gramatykę będącą w słabej postaci normalnej Chomsky'ego:

$S \to SXSS + b$
$X \to a$

Rozpatrzmy rodziny $\mathcal{L}$ i $\mathcal{R}$.

$$\begin{array}{ll}
L(a,S) = \varnothing & L(b,S) = (XSS)^\star = L_0 \\
L(a,X) = \{\epsilon\} = E & L(b,X) = \varnothing \\
R(a,S) = \varnothing & R(b,S) = (SXS)^\star = L_1 \\
R(a,X) = \{\epsilon\} = E & R(b,X) = \varnothing
\end{array}$$

$\mathcal{L} = \{\varnothing, L_0, E\} \qquad \mathcal{R} = \{\varnothing, L_1, E\}$

Obliczmy domknięcie $\mathcal{L} \bigcup \mathcal{R}$ względem lewostronnego oraz prawostronnego ilorazu.

$$\begin{array}{ll}
X^{-1}L_0 = SS(XSS)^\star = L_2 & L_0S^{-1} = (XSS)^\star XS = L_3 \\
S^{-1}L_1 = XS(SXS)^\star = L_3 & L_1S^{-1} = (SXS)^\star SX = L_4 \\
S^{-1}L_2 = S(XSS)^\star = L_5 & L_2S^{-1} = (SSX)^\star S = L_6 \\
X^{-1}L_3 = S(SXS)^\star = L_6 & L_3S^{-1} = (XSS)^\star X = L_7 \\
S^{-1}L_4 = X(SSX)^\star = L_7 & L_4X^{-1} = (SXS)^\star S = L_5 \\
S^{-1}L_5 = (XSS)^\star = L_0 & L_5S^{-1} = (SXS)^\star = L_1 \\
S^{-1}L_6 = (SXS)^\star = L_1 & L_6S^{-1} = (SSX)^\star = L_8 \\
X^{-1}L_7 = (SSX)^\star = L_8 & L_7X^{-1} = (XSS)^\star = L_0 \\
S^{-1}L_8 = SX(SSX)^\star = L_4 & L_8X^{-1} = (SSX)^\star SS = L_2
\end{array}$$

Równoważne gramatyki wyglądają następująco:

$$
\begin{array}{ll}
S \to bL_0 + bE & S \to L_1 b + Eb \\
X \to aE & X \to Ea \\
L_0 \to XL_2 + \epsilon & L_0 \to L_3 S + \epsilon \\
L_1 \to SL_3 + \epsilon & L_1 \to L_4 S + \epsilon \\
L_2 \to SL_5 & L_2 \to L_6 S \\
L_3 \to XL_6 & L_3 \to L_7 S \\
L_4 \to SL_7 & L_4 \to L_5 X \\
L_5 \to SL_0 & L_5 \to L_1 S \\
L_6 \to SL_1 & L_6 \to L_8 S \\
L_7 \to XL_8 & L_7 \to L_0 X \\
L_8 \to SL_4 + \epsilon & L_8 \to L_2 X + \epsilon \\
E \to \epsilon & E \to \epsilon
\end{array}
$$

Zastępując X przez aE i S przez $bL_0 + bE$ w lewej gramatyce oraz X przez Ea i S przez $L_1 b + Eb$ w prawej gramatyce otrzymujemy:

$$
\begin{array}{ll}
S \to bL_0 + bE & S \to L_1 b + Eb \\
X \to aE & X \to Ea \\
L_0 \to aEL_2 + \epsilon & L_0 \to L_3 L_1 b + L_3 Eb + \epsilon \\
L_1 \to bL_0 L_3 + bEL_3 + \epsilon & L_1 \to L_4 L_1 b + L_4 Eb + \epsilon \\
L_2 \to bL_0 L_5 + bEL_5 & L_2 \to L_6 L_1 b + L_6 Eb \\
L_3 \to aEL_6 & L_3 \to L_7 L_1 b + L_7 Eb \\
L_4 \to bL_0 L_7 + bEL_7 & L_4 \to L_5 Ea \\
L_5 \to bL_0 L_0 + bEL_0 & L_5 \to L_1 L_1 b + L_1 Eb \\
L_6 \to bL_0 L_1 + bEL_1 & L_6 \to L_8 L_1 b + L_8 Eb \\
L_7 \to aEL_8 & L_7 \to L_0 Ea \\
L_8 \to bL_0 L_4 + bEL_4 + \epsilon & L_8 \to L_2 Ea + \epsilon \\
E \to \epsilon & E \to \epsilon
\end{array}
$$

Po usunięciu ϵ-produkcji mamy:

$$
\begin{array}{ll}
S \to bL_0 + b & S \to L_1 b + b \\
X \to a & X \to a \\
L_0 \to aL_2 & L_0 \to L_3 L_1 b + L_3 b \\
L_1 \to bL_0 L_3 + bL_3 & L_1 \to L_4 L_1 b + L_4 b \\
L_2 \to bL_0 L_5 + bL_5 & L_2 \to L_6 L_1 b + L_6 b \\
L_3 \to aL_6 & L_3 \to L_7 L_1 b + L_7 b \\
L_4 \to bL_0 L_7 + bL_7 & L_4 \to L_5 a \\
L_5 \to bL_0 L_0 + bL_0 + b & L_5 \to L_1 L_1 b + L_1 b + b \\
L_6 \to bL_0 L_1 + bL_0 + bL_1 + b & L_6 \to L_8 L_1 b + L_8 b + L_1 b + b \\
L_7 \to aL_8 + a & L_7 \to L_0 a + a \\
L_8 \to bL_0 L_4 + bL_4 & L_8 \to L_2 a
\end{array}
$$

Otrzymujemy gramatykę w sześciennej podwójnej postaci normalnej Greibach.

$$
\begin{aligned}
S &\to bL_3L_1b + bL_3b + b \\
X &\to a \\
L_0 &\to aL_6L_1b + aL_6b \\
L_1 &\to bL_0L_7L_1b + bL_0L_7b + bL_7L_1b + bL_7b \\
L_2 &\to bL_0L_1L_1b + bL_0L_1b + bL_0b + bL_1L_1b + bL_1b + bb \\
L_3 &\to aL_8L_1b + aL_8b + aL_1b + ab \\
L_4 &\to bL_0L_0a + bL_0a + ba \\
L_5 &\to bL_0L_3L_1b + bL_0L_3b + bL_3L_1b + bL_3b + b \\
L_6 &\to bL_0L_4L_1b + bL_0L_4b + bL_3L_1b + bL_3b + bL_4L_1b + bL_4b + b \\
L_7 &\to aL_2a + a \\
L_8 &\to bL_0L_5a + bL_5a
\end{aligned}
$$

Wprowadzamy nowe nieterminale oraz produkcje celem sprowadzenia gramatyki do kwadratowej podwójnej postaci normalnej Greibach.

$$
\begin{aligned}
S &\to bL_3L_1b + bL_3b + b \\
X &\to a \\
L_0 &\to aL_6L_1b + aL_6b \\
L_1 &\to b\langle L_0L_7\rangle L_1b + bL_0L_7b + bL_7L_1b + bL_7b \\
L_2 &\to b\langle L_0L_1\rangle L_1b + bL_0L_1b + bL_0b + bL_1L_1b + bL_1b + bb \\
L_3 &\to aL_8L_1b + aL_8b + aL_1b + ab \\
L_4 &\to bL_0L_0a + bL_0a + ba \\
L_5 &\to b\langle L_0L_3\rangle L_1b + bL_0L_3b + bL_3L_1b + bL_3b + b \\
L_6 &\to b\langle L_0L_4\rangle L_1b + bL_0L_4b + bL_3L_1b + bL_3b + bL_4L_1b + bL_4b + b \\
L_7 &\to aL_2a + a \\
L_8 &\to bL_0L_5a + bL_5a \\
\langle L_0L_1\rangle &\to aL_2\langle L_4L_1\rangle b + aL_2L_4b \\
\langle L_0L_3\rangle &\to a\langle L_2L_7\rangle L_1b + aL_2L_7b \\
\langle L_0L_4\rangle &\to aL_2L_5a \\
\langle L_0L_7\rangle &\to aL_2L_0a + aL_2a \\
\langle L_2L_7\rangle &\to b\langle L_0L_5\rangle L_0a + bL_0L_5a + bL_5L_0a + bL_5a \\
\langle L_0L_5\rangle &\to a\langle L_2L_1\rangle L_1b + aL_2L_1b + aL_2b \\
\langle L_2L_1\rangle &\to b\langle L_0L_5\rangle\langle L_4L_1\rangle b + b\langle L_0L_5\rangle L_4b + bL_5\langle L_4L_1\rangle b + bL_5L_4b \\
\langle L_4L_1\rangle &\to b\langle L_0L_7\rangle\langle L_4L_1\rangle b + b\langle L_0L_7\rangle L_4b + bL_7\langle L_4L_1\rangle b + bL_7L_4b
\end{aligned}
$$

5 Metoda Bluma-Kocha

W metodzie Bluma-Kocha zakłada się, że symbol początkowy gramatyki nie występuje po prawej stronie żadnej produkcji. Warunek ten łatwo spełnić w dowolnej gramatyce wprowadzając w razie potrzeby nowy symbol początkowy gramatyki oraz produkcję $S' \to S$, gdzie S jest poprzednim symbolem początkowym gramatyki. Warunek ten będziemy nazywać warunkiem Bluma-Kocha.
Wynikiem działania algorytmu jest gramatyka w tzw. postaci normalnej z nagłówkami terminalnymi (ang. normal form with terminal heads). Podobnie jak to czyniliśmy przy okazji metody Rosenkrantz'a, gramatykę taką można prosto sprowadzić do postaci normalnej Greibach wprowadzając nowe zmienne pomocnicze C_a dla każdego terminala $a \in T$ oraz produkcje postaci $C_a \to a$ i zamieniając każdy nie skrajnie lewy terminal a występujący w prawych stronach produkcji przez nieterminal C_a.
Algorytm powoduje zwiększenie rozmiaru gramatyki rzędu $O(|G|^3)$, gdy gramatyka nie zawiera produkcji łańcuchowych oraz rzędu $O(|G|^4)$, gdy gramatyka zawiera produkcje łańcuchowe.

5.1 Sprowadzanie gramatyki do postaci normalnej Greibach

Algorytm.

1. Dla każdego nieterminala $B \in N - \{S\}$ utwórz gramatykę $G_B = \langle N_B, N \cup T, P_B, S_B \rangle$ gdzie $N_B = \{ A_B \mid A \in N \} \cup N$.
   Zauważmy, że w gramatyce G_B zbiorem terminali jest zbiór terminali i nieterminali gramatyki G. Niech $N(A)$ oznacza zbiór nieterminali otrzymanych z A przy użyciu produkcji łańcuchowych. t.j. $N(A) = \{ C \in N \mid A \Rightarrow^* C \}$. Produkcje P_B do gramatyki G_B dodawane są następująco:

Dodawane produkcje	Warunek
$S_B \to a\gamma$	$C \to a\gamma \in P$, $C \in N(B)$, $a \in T$, $\gamma \in (N \cup T)^\star$
$S_B \to a\gamma C_B$	$C \to a\gamma \in P$, $a \in T$, $\gamma \in (N \cup T)^\star$
$C_B \to \alpha D_B$	$D \to C\alpha \in P$, $D \in N - \{S\}$, $C \in N$, $\alpha \in (N \cup T)^\star$
$C_B \to \alpha$	$D \to C\alpha \in P$, $D \in N(B)$, $C \in N$, $\alpha \in (N \cup T)^+$

   Dodaj P_B do P.
2. Dla dowolnych symboli nieterminalnych $B, E \in N - \{S\}$
   - każdą produkcję $A \to B\alpha$ zastąp przez $A \to S_B\alpha$
   - każdą produkcję $A_E \to B\alpha$ zastąp przez $A_E \to S_B\alpha$
3. Dla dowolnych symboli nieterminalnych $B, E \in N - \{S\}$
   - każdą produkcję $A \to S_B\alpha$ zastąp przez $\{ A \to a\gamma\alpha \mid S_B \to a\gamma \in P_B \}$
   - każdą produkcję $A_E \to S_B\alpha$ zastąp przez $\{ A_E \to a\gamma\alpha \mid S_B \to a\gamma \in P_B \}$
4. Dla każdego nieterminala $B \in N - \{S\}$ usuń $\{S_B \to \alpha \mid S_B \to \alpha \in P_B\}$ ze zbioru produkcji P.
5. Jeżeli w gramatyce istnieją produkcje łańcuchowe postaci $D_E \to C_E$ to należy je wyeliminować w standardowy sposób.

Przykład 5.1 Rozważmy gramatykę

$X \to XaY \mid b$
$Y \to YbX \mid a$

Wprowadzając nowy symbol początkowy S tworzymy gramatykę rozszerzoną $G = \langle \{S, X, Y\}, \{a, b\}, P, S \rangle$, gdzie P:

$S \to X$
$X \to XaY \mid b$
$Y \to YbX \mid a$

Zgodnie z punktem 1. algorytmu tworzymy nowe gramatyki:

Gramatyka $G_X = \langle \{S_X, X_X, Y_X\}, \{a, b, S, X, Y\}, P_X, S_X \rangle$:
$S_X \to b \mid bX_X \mid aY_X$
$X_X \to aY \mid aYX_X$
$Y_X \to bXY_X$

Gramatyka $G_Y = \langle \{S_Y, X_Y, Y_Y\}, \{a, b, S, X, Y\}, P_Y, S_Y \rangle$:
$S_Y \to a \mid aY_Y \mid bX_Y$
$X_Y \to aYX_Y$
$Y_Y \to bX \mid bXY_Y$

Choć nie jest to zaznaczone w algorytmie warto w tej fazie pozbyć się symboli nieużytecznych (tj. nieterminali Y_X, X_Y). Po ich usunięciu otrzymujemy:

Gramatyka G_X:
$S_X \to b \mid bX_X$
$X_X \to aY \mid aYX_X$

Gramatyka G_Y:
$S_Y \to a \mid aY_Y$
$Y_Y \to bX \mid bXY_Y$

Gramatyki G_X i G_Y generują odpowiednio języki $L_X = \{\, b(aY)^k \mid k \geq 0 \,\}$ oraz $L_Y = \{\, a(bX)^k \mid k \geq 0 \,\}$, co zgadza sie z przedstawioną definicją tych gramatyk.
Po dodaniu produkcji gramatyk G_X i G_Y gramatyka G ma postać:
$S \to X$
$X \to XaY \mid b$
$Y \to YbX \mid a$
$S_X \to b \mid bX_X$
$X_X \to aY \mid aYX_X$
$S_Y \to a \mid aY_Y$
$Y_Y \to bX \mid bXY_Y$

W kroku 2. zastępujemy produkcje:

$S \to X$
$X \to XaY$
$Y \to YbX$

przez

$S \to S_X$
$X \to S_XaY$
$Y \to S_YbX$

a z kolei w kroku 3. powyższe produkcje zastępujemy przez

$S \to b \mid bX_X$
$X \to baY \mid bX_XaY$
$Y \to abX \mid aY_YbX$

Wszystkie produkcje zaczynające się nieterminalami S_X i S_Y po lewej stronie są usuwane.
Ostatecznie otrzymujemy gramatykę w postaci normalnej z prawymi stronami zaczynającymi się od terminali:

$S \to b \mid bX_X$
$X \to b \mid baY \mid bX_XaY$
$Y \to a \mid abX \mid aY_YbX$
$X_X \to aY \mid aYX_X$
$Y_Y \to bX \mid bXY_Y$

Gramatyka nie zawiera produkcji łańcuchowych, i łatwo otrzymać z niej postać normalną Greibach.

Przykład 5.2 Rozważmy gramatykę

$S \to SXSS \mid b$
$X \to a$

Wprowadzamy nowy symbol początkowy S' celem spełnienia warunku Bluma-Kocha. Otrzymujemy gramatykę $G = \langle \{S', S, X\}, \{a, b\}, P, S' \rangle$ gdzie P:

$S' \to S$
$S \to SXSS \mid b$
$X \to a$

Tworzymy gramatyki G_S i G_X.

Gramatyka $G_S = \langle \{S'_S, S_S, X_S\}, \{S', S, X, a, b\}, P_S, S'_S \rangle$:

$S'_S \to b \mid bS_S \mid aX_S$
$S_S \to XSS \mid XSSS_S$

Gramatyka $G_X = \langle\{S'_X, S_X, X_X\}, \{S', S, X, a, b\}, P_X, S'_X\rangle$:

$S'_X \to a \mid aX_X \mid bS_X$
$S_X \to XSSS_X$

Modyfikujemy gramatyki G_S, G_X usuwając symbole bezużyteczne:

Gramatyka G_S:

$S'_S \to b \mid bS_S$
$S_S \to XSS \mid XSSS_S$

Gramatyka G_X:

$S'_X \to a$

Gramatyka G ma na tym etapie postać:
$S' \to S$
$S \to SXSS \mid b$
$X \to a$
$S'_S \to b \mid bS_S$
$S_S \to XSS \mid XSSS_S$
$S'_X \to a$

Dokonujemy zastąpień wg 2. i 3. kroku algorytmu:

Usuwana produkcja	I zastąpienie	II zastąpienie
$S' \to S$	$S' \to S'_S$	$S' \to b \mid bS_S$
$S \to SXSS$	$S \to S'_SXSS$	$S \to bXSS \mid bS_SXSS$
$S_S \to XSSS_S$	$S_S \to S'_XSSS_S$	$S_S \to aSSS_S$
$S_S \to XSS$	$S_S \to S'_XSS$	$S_S \to aSS$

Po usunięciu produkcji typu $S'_B \to \alpha, B \in N - \{S'\}$ dostajemy:

$S' \to b \mid bS_S$
$S \to b \mid bXSS \mid bS_SXSS$
$S_S \to aSS \mid aSSS_S$
$X \to a$

Otrzymana gramatyka nie zawiera żadnych produkcji łańcuchowych i łatwo już otrzymać z niej postać normalną Greibach.

5.2 Sprowadzanie gramatyki do kwadratowej postaci normalnej Greibach (2-GNF)

Podobnie jak poprzednia konstrukcja algorytm bazuje na konstrukcji języków typu L_B dla $B \in N - \{S\}$. Są to języki słów $x \in T(N - \{S\})$ w alfabecie terminali i nieterminali $N \cup T$.
W odróżnieniu od podstawowej wersji algorytmu Bluma-Kocha algorytm sprowadzania do postaci kwadratowej 2-GNF wymaga na wejściu gramatyki w postaci normalnej Chomsky'ego oraz bazuje na konstrukcjach automatów skończonych opisujących języki L_B oraz L_B^R. Powoduje on wzrost rozmiaru gramatyki rzędu $O(|G|^3)$.
Istnieją pewne modyfikacje opisywanego algorytmu, w których konstrukcja gramatyki w postaci normalnej Greibach przebiega poprzez gramatykę w rozszerzonej postaci normalnej Chomsky'ego oraz rozszerzonej postaci normalnej Greibach (dopuszczalne są produkcje łańcuchowe). Wtedy cała konstrukcja powoduje wzrost rozmiaru gramatyki rzędu $O(|G|^4)$. Gdyby zaś konstrukcja przebiegała torem $G \longrightarrow$

gramatyka w postaci Chomsky'ego $\longrightarrow$ gramatyka w postaci Greibach (a więc bez produkcji łańcuchowych), powodowałaby wzrost rozmiaru gramatyki rzędu $O(|G|^6)$.
Poniżej omawiana jest podstawowa metoda, od gramatyki w postaci normalnej Chomsky'ego do gramatyki w postaci normalnej Greibach.

Algorytm.

1. Dla każdego $B \in N - \{S\}$ skonstruuj automat $M_B = (Q, V, \delta, B_B, \mathcal{S}_B)$ opisujący język $L_B^R = L(M_B)$, gdzie:
   $Q_B = \{\, A_B \mid A \in N \,\} \cup \{\mathcal{S}_B\}$
   $\delta(C_B, E) = \{\, D_B \mid C \to DE \in P \,\}$
   $\delta(C_B, a) = \begin{cases} \{\mathcal{S}_B\} & \text{jeśli } C \to a \in P \\ \varnothing & \text{w przeciwnym wypadku} \end{cases}$
   dla każdego $C_B \in Q$, $E \in N - \{S\}$, $a \in T$

   Należy podkreslic różnicę między stanem S_B a stanem $\mathcal{S}_B$. W przykładach zostanie pokazane, że niepotrzebne jest branie pod uwagę stanu S_B, gdyż nie jest on wykorzystywany do późniejszych konstrukcji gramatyk G_B.

2. Dla każdego $B \in N - \{S\}$ skonstruuj automat $M'_B = (Q, V, \delta', S_B, B_B)$ opisujący język $L_B = L(M'_B)$, gdzie:
   $Q_B = \{\, A_B \mid A \in N \,\} \cup \{\mathcal{S}_B\}$
   $\delta'(\mathcal{S}_B, a) = \{\, C_B \mid \{\mathcal{S}_B\} = \delta(C_B, a) \,\}$
   $\delta'(D_B, E) = \{\, C_B \mid D_B \in \delta(C_B, E) \,\}$
   dla każdego $D_B \in Q$, $E \in N - \{S\}$, $a \in T$

   Przy konstruowaniu automatu M'_B korzystamy z faktu, że aby z języka L_B^R otrzymać L_B wystarczy w automacie M_B odwrócić wszystkie przejścia, zamienić stan końcowy automatu M_B na początkowy w M'_B, zaś początkowy w M_B na końcowy w M'_B.

3. Dla każdego $B \in N - \{S\}$ zdefiniuj gramatykę $G'_B = \langle V_B, V, P'_B, \mathcal{S}_B \rangle$ gdzie $V_B = \{\, A_B \mid A \in N - \{S\} \,\} \cup \{\mathcal{S}_B\} \cup V$.
   Produkcje P'_B do gramatyki G'_B dodawane są następująco:

Dodawane produkcje	Warunek
$\mathcal{S}_B \to aC_B$	$C_B \in \delta'(\mathcal{S}_B, a)$ oraz $C_B \neq B_B$ lub $\delta'(\{B_B\} \times N - \{S\}) \neq \varnothing$
$\mathcal{S}_B \to a$	$B_B \in \delta'(\mathcal{S}_B, a)$
$D_B \to EC_B$	$C_B \in \delta'(D_B, E)$ oraz $C_B \neq B_B$ lub $\delta'(\{B_B\} \times N - \{S\}) \neq \varnothing$
$D_B \to E$	$B_B \in \delta'(D_B, E)$

   dla każdego $D_B \in V'_B,\ E \in N - \{S\},\ a \in \Sigma$, gdzie podobno? $V'_B = V_B - V,\ V = N \cup T$

   Następnie dla każdego $B \in N - \{S\}$ z gramatyki G'_B usuń symbole nieużyteczne.

4. Dla każdego $B \in N - \{S\}$ skonstruuj gramatykę $G_B = \langle V_B, T, P_B, \mathcal{S}_B \rangle$.
   Produkcje P_B do gramatyki G_B dodawane są następująco:

Dodawane produkcje	Warunek
$\mathcal{S}_B \to aC_B$	$\mathcal{S}_B \to aC_B \in P'_B$
$\mathcal{S}_B \to a$	$\mathcal{S}_B \to a \in P'_B$
$D_B \to \alpha C_B$	$D_B \to EC_B \in P'_B$, α jest alternatywą $\mathcal{S}_E$ w G'_E
$D_B \to \alpha$	$D_B \to E \in P'_B$, α jest alternatywą $\mathcal{S}_E$ w G'_E

5. Skonstruuj gramatykę $G' = \langle V', T, P', S \rangle$ gdzie $V' = V \cup \{\, N_B \mid N - \{S\} \,\}$.
   Produkcje do gramatyki G' dodawane są następująco:

Dodawane produkcje	Warunek
$A \to a$	$A \to a \in P$
$A \to a\gamma C$	$A \to BC \in P$, $a\gamma$ jest alternatywą $\mathcal{S}_B$ w G_B
$\bigcup_{B\in N-\{S\}} P_B - \{\, S_B \to \alpha \mid \alpha \in TN_B^\star \,\}$	

Przykład 5.3 Ponownie rozważmy gramatykę:

$S \to SXSS \mid b$
$X \to a$

Celem spełnienia warunku Bluma-Kocha wprowadzamy nowy symbol początkowy S':

$S' \to S$
$S \to SXSS \mid b$
$X \to a$

Następnie przekształcamy gramatykę do postaci normalnej Chomsky'ego:

$S' \to SA \mid b$
$S \to SA \mid b$
$A \to XB$
$B \to SS$
$X \to a$

Oznaczmy:
$Q_S = \{S'_S, A_S, B_S, X_S, S_S, \mathcal{S}'_S\}$
$Q_A = \{S'_A, A_A, B_A, X_A, S_A, \mathcal{S}'_A\}$
$Q_B = \{S'_B, A_B, B_B, X_B, S_B, \mathcal{S}'_B\}$
$Q_X = \{S'_X, A_X, B_X, X_X, S_X, \mathcal{S}'_X\}$

Konstruujemy automaty $M_B,\ B \in N - \{S'\}$:

$M_S = (Q_S, V, \delta_S, S_S, \mathcal{S}'_S)$	$M_A = (Q_A, V, \delta_A, A_A, \mathcal{S}'_A)$	$M_B = (Q_B, V, \delta_B, B_B, \mathcal{S}'_B)$	$M_X = (Q_X, V, \delta_X, X_X, \mathcal{S}'_X)$
$\delta(S'_S, A) = S_S$	$\delta(S'_A, A) = S_A$	$\delta(S'_B, A) = S_B$	$\delta(S'_X, A) = S_X$
$\delta(S_S, A) = S_S$	$\delta(S_A, A) = S_A$	$\delta(S_B, A) = S_B$	$\delta(S_X, A) = S_X$
$\delta(A_S, B) = X_S$	$\delta(A_A, B) = X_A$	$\delta(A_B, B) = X_B$	$\delta(A_X, B) = X_X$
$\delta(B_S, S) = S_S$	$\delta(B_A, S) = S_A$	$\delta(B_B, S) = S_B$	$\delta(B_X, S) = S_X$
$\delta(S'_S, b) = \mathcal{S}'_S$	$\delta(S'_A, b) = \mathcal{S}'_A$	$\delta(S'_B, b) = \mathcal{S}'_B$	$\delta(S'_X, b) = \mathcal{S}'_X$
$\delta(S_S, b) = \mathcal{S}'_S$	$\delta(S_A, b) = \mathcal{S}'_A$	$\delta(S_B, b) = \mathcal{S}'_B$	$\delta(S_X, b) = \mathcal{S}'_X$
$\delta(X_S, a) = \mathcal{S}'_S$	$\delta(X_A, a) = \mathcal{S}'_A$	$\delta(X_B, a) = \mathcal{S}'_B$	$\delta(X_X, a) = \mathcal{S}'_X$

oraz automaty $M'_B,\ B \in N - \{S'\}$:

$M'_S = (Q_S, V, \delta'_S, \mathcal{S}'_S, S_S)$	$M'_A = (Q_A, V, \delta'_A, \mathcal{S}'_A, A_A)$	$M'_B = (Q_B, V, \delta'_B, \mathcal{S}'_B, B_B)$	$M'_X = (Q_X, V, \delta'_X, \mathcal{S}'_X, X_X)$
$\delta'(S_S, A) = \{S'_S, S_S\}$	$\delta'(S_A, A) = \{S'_A, S_A\}$	$\delta'(S_B, A) = \{S'_B, S_B\}$	$\delta'(S_X, A) = \{S'_X, S_X\}$
$\delta'(X_S, B) = A_S$	$\delta'(X_A, B) = A_A$	$\delta'(X_B, B) = A_B$	$\delta'(X_X, B) = A_X$
$\delta'(S_S, S) = B_S$	$\delta'(S_A, S) = B_A$	$\delta'(S_B, S) = B_B$	$\delta'(S_X, S) = B_X$
$\delta'(\mathcal{S}'_S, b) = \{S'_S, S_S\}$	$\delta'(\mathcal{S}'_A, b) = \{S'_A, S_A\}$	$\delta'(\mathcal{S}'_B, b) = \{S'_B, S_B\}$	$\delta'(\mathcal{S}'_X, b) = \{S'_X, S_X\}$
$\delta'(\mathcal{S}'_S, a) = X_S$	$\delta'(\mathcal{S}'_A, a) = X_A$	$\delta'(\mathcal{S}'_B, a) = X_B$	$\delta'(\mathcal{S}'_X, a) = X_X$

Powyższe automaty służą do skonstruowania następujących gramatyk:

Gramatyka G'_S:	Gramatyka G'_A:	Gramatyka G'_B:	Gramatyka G'_X:
$\mathcal{S}'_S \to bS_S \mid aX_S \mid b$ $S_S \to AS_S \mid SB_S \mid A$ $X_S \to BA_S$	$\mathcal{S}'_A \to bS_A \mid aX_A$ $S_A \to AS_A \mid SB_A$ $X_A \to B$	$\mathcal{S}'_B \to bS_B \mid aX_B$ $S_B \to AS_B \mid S$ $X_B \to BA_B$	$\mathcal{S}'_X \to bS_X \mid aX_X \mid a$ $S_X \to AS_X \mid SB_X$ $X_X \to BA_X$
Po usunięciu symboli nieużytecznych:			
$\mathcal{S}'_S \to bS_S \mid b$ $S_S \to AS_S \mid A$	$\mathcal{S}'_A \to aX_A$ $X_A \to B$	$\mathcal{S}'_B \to bS_B$ $S_B \to AS_B \mid S$	$\mathcal{S}'_X \to a$
$L(G'_S) = \{\, bA^k \mid k \geq 0 \,\}$	$L(G'_A) = \{aB\}$	$L(G'_B) = \{\, bA^kY \mid k \geq 0 \,\}$	$L(G'_X) = \{a\}$

Gramatyka G_S:

$\mathcal{S}'_S \to bS_S \mid b$
$S_S \to aX_AS_S \mid aX_A$

Gramatyka G_A:

$\mathcal{S}'_A \to aX_A$
$X_A \to bS_B$

Gramatyka G_B:

$\mathcal{S}'_B \to bS_B$
$S_B \to aX_AS_B \mid bS_S \mid b$

Gramatyka G_X:

$\mathcal{S}'_X \to a$

Gramatyka G':

$S' \to b$
$S \to b$
$X \to a$
$S' \to bS_SA \mid bA$
$S \to bS_SA \mid bA$
$A \to aB$
$B \to bS_SS \mid bS$
$S_S \to aX_AS_S$
$S_S \to aX_A$
$X_A \to bS_B$
$S_B \to aX_AS_B$
$S_B \to bS_S \mid b$

Po usunięciu symboli bezużytecznych i nieosiągalnych dostajemy gramatykę zawierającą zbiór nieterminali $N = \{\, S', S, A, B, S_S, S_B, X_A \,\}$ i produkcje:

$S' \to b \mid bA \mid bS_SA$
$S \to b \mid bA \mid bS_SA$
$A \to aB$
$B \to bS \mid bS_SS$
$S_S \to aX_A \mid aX_AS_S$
$X_A \to bS_B$
$S_B \to b \mid bS_S \mid aX_AS_B$

Przykład 5.4 Rozważmy gramatykę

$X \to XaY \mid b$
$Y \to YbX \mid a$

Tak jak w poprzednim przykładzie wprowadzamy nowy symbol początkowy S:

$S \to X$
$X \to XaY \mid b$
$Y \to YbX \mid a$

a następnie przekształcamy gramatykę do postaci normalnej Chomsky'ego:

$S \to XC \mid b$
$X \to XC \mid b$
$C \to AY$
$Y \to YD \mid a$
$D \to BX$
$A \to a$
$B \to b$

Krok 1. Konstruujemy automaty $M_B,\ B \in N - \{S'\}$.

Oznaczmy:

$Q_X = \{S_X, A_X, B_X, C_X, D_X, X_X, Y_X, \mathcal{S}_X\}$
$Q_A = \{S_A, A_A, B_A, C_A, D_A, X_A, Y_A, \mathcal{S}_A\}$
$Q_B = \{S_B, A_B, B_B, C_B, D_B, X_B, Y_B, \mathcal{S}_B\}$
$Q_C = \{S_C, A_C, B_C, C_C, D_C, X_C, Y_C, \mathcal{S}_C\}$
$Q_D = \{S_D, A_D, B_D, C_D, D_D, X_D, Y_D, \mathcal{S}_D\}$
$Q_Y = \{S_Y, A_Y, B_Y, C_Y, D_Y, X_Y, Y_Y, \mathcal{S}_Y\}$

$M_X = (Q_X, N \cup T, \delta_X, X_X, \mathcal{S}_X)$	$M_A = (Q_A, N \cup T, \delta_A, A_A, \mathcal{S}_A)$	$M_B = (Q_B, N \cup T, \delta_B, B_B, \mathcal{S}_B)$
$\delta(A_X, a) = \{\mathcal{S}_X\}$	$\delta(A_A, a) = \{\mathcal{S}_A\}$	$\delta(A_B, a) = \{\mathcal{S}_B\}$
$\delta(B_X, b) = \{\mathcal{S}_X\}$	$\delta(B_A, b) = \{\mathcal{S}_A\}$	$\delta(B_B, b) = \{\mathcal{S}_B\}$
$\delta(C_X, Y) = \{A_X\}$	$\delta(C_A, Y) = \{A_A\}$	$\delta(C_B, Y) = \{A_B\}$
$\delta(D_X, X) = \{B_X\}$	$\delta(D_A, X) = \{B_A\}$	$\delta(D_B, X) = \{B_B\}$
$\delta(X_X, b) = \{\mathcal{S}_X\}$	$\delta(X_A, b) = \{\mathcal{S}_A\}$	$\delta(X_B, b) = \{\mathcal{S}_B\}$
$\delta(X_X, C) = \{X_X\}$	$\delta(X_A, C) = \{X_A\}$	$\delta(X_B, C) = \{X_B\}$
$\delta(Y_X, a) = \{\mathcal{S}_X\}$	$\delta(Y_A, a) = \{\mathcal{S}_A\}$	$\delta(Y_B, a) = \{\mathcal{S}_B\}$
$\delta(Y_X, D) = \{Y_X\}$	$\delta(Y_A, D) = \{Y_A\}$	$\delta(Y_B, D) = \{Y_B\}$
$\delta(S_X, b) = \{\mathcal{S}_X\}$	$\delta(S_A, b) = \{\mathcal{S}_A\}$	$\delta(S_B, b) = \{\mathcal{S}_B\}$
$\delta(S_X, C) = \{X_X\}$	$\delta(S_A, C) = \{X_A\}$	$\delta(S_B, C) = \{X_B\}$

$M_C = (Q_C, N \cup T, \delta_C, C_C, \mathcal{S}_C)$	$M_D = (Q_D, N \cup T, \delta_D, D_D, \mathcal{S}_D)$	$M_Y = (Q_Y, N \cup T, \delta_Y, Y_Y, \mathcal{S}_Y)$
$\delta(A_C, a) = \{\mathcal{S}_C\}$	$\delta(A_D, a) = \{\mathcal{S}_D\}$	$\delta(A_Y, a) = \{\mathcal{S}_Y\}$
$\delta(B_C, b) = \{\mathcal{S}_C\}$	$\delta(B_D, b) = \{\mathcal{S}_D\}$	$\delta(B_Y, b) = \{\mathcal{S}_Y\}$
$\delta(C_C, Y) = \{A_C\}$	$\delta(C_D, Y) = \{A_D\}$	$\delta(C_Y, Y) = \{A_Y\}$
$\delta(D_C, X) = \{B_C\}$	$\delta(D_D, X) = \{B_D\}$	$\delta(D_Y, X) = \{B_Y\}$
$\delta(X_C, b) = \{\mathcal{S}_C\}$	$\delta(X_D, b) = \{\mathcal{S}_D\}$	$\delta(X_Y, b) = \{\mathcal{S}_Y\}$
$\delta(X_C, C) = \{X_C\}$	$\delta(X_D, C) = \{X_D\}$	$\delta(X_Y, C) = \{X_Y\}$
$\delta(Y_C, a) = \{\mathcal{S}_C\}$	$\delta(Y_D, a) = \{\mathcal{S}_D\}$	$\delta(Y_Y, a) = \{\mathcal{S}_Y\}$
$\delta(Y_C, D) = \{Y_C\}$	$\delta(Y_D, D) = \{Y_D\}$	$\delta(Y_Y, D) = \{Y_Y\}$
$\delta(S_C, b) = \{\mathcal{S}_C\}$	$\delta(S_D, b) = \{\mathcal{S}_D\}$	$\delta(S_Y, b) = \{\mathcal{S}_Y\}$
$\delta(S_C, C) = \{X_C\}$	$\delta(S_D, C) = \{X_D\}$	$\delta(S_Y, C) = \{X_Y\}$

oraz automaty $M'_B,\ B \in N - \{S'\}$:

$M'_X = (Q_X, N \cup T, \delta'_X, \mathcal{S}_X, X_X)$	$M'_A = (Q_A, N \cup T, \delta'_A, \mathcal{S}_A, A_A)$	$M'_B = (Q_B, N \cup T, \delta'_B, \mathcal{S}_B, B_B)$
$\delta'(A_X, Y) = \{C_X\}$ $\delta'(B_X, X) = \{D_X\}$ $\delta'(X_X, C) = \{X_X, S_X\}$ $\delta'(Y_X, D) = \{Y_X\}$ $\delta'(\mathcal{S}_X, b) = \{X_X, S_X, B_X\}$ $\delta'(\mathcal{S}_X, a) = \{Y_X, A_X\}$	$\delta'(A_A, Y) = \{C_A\}$ $\delta'(B_A, X) = \{D_A\}$ $\delta'(X_A, C) = \{X_A, S_A\}$ $\delta'(Y_A, D) = \{Y_A\}$ $\delta'(\mathcal{S}_A, b) = \{X_A, S_A, B_A\}$ $\delta'(\mathcal{S}_A, a) = \{Y_A, A_A\}$	$\delta'(A_B, Y) = \{C_B\}$ $\delta'(B_B, X) = \{D_B\}$ $\delta'(X_B, C) = \{X_B, S_B\}$ $\delta'(Y_B, D) = \{Y_B\}$ $\delta'(\mathcal{S}_B, b) = \{X_B, S_B, B_B\}$ $\delta'(\mathcal{S}_B, a) = \{Y_B, A_B\}$

$M'_C = (Q_C, N \cup T, \delta'_C, \mathcal{S}_C, C_C)$	$M'_D = (Q_D, N \cup T, \delta'_D, \mathcal{S}_D, D_D)$	$M'_Y = (Q_Y, N \cup T, \delta'_Y, \mathcal{S}_Y, Y_Y)$
$\delta'(A_C, Y) = \{C_C\}$ $\delta'(B_C, X) = \{D_C\}$ $\delta'(X_C, C) = \{X_C, S_C\}$ $\delta'(Y_C, D) = \{Y_C\}$ $\delta'(\mathcal{S}_C, b) = \{X_C, S_C, B_C\}$ $\delta'(\mathcal{S}_C, a) = \{Y_C, A_C\}$	$\delta'(A_D, Y) = \{C_D\}$ $\delta'(B_D, X) = \{D_D\}$ $\delta'(X_D, C) = \{X_D, S_D\}$ $\delta'(Y_D, D) = \{Y_D\}$ $\delta'(\mathcal{S}_D, b) = \{X_D, S_D, B_D\}$ $\delta'(\mathcal{S}_D, a) = \{Y_D, A_D\}$	$\delta'(A_Y, Y) = \{C_Y\}$ $\delta'(B_Y, X) = \{D_Y\}$ $\delta'(X_Y, C) = \{X_Y, S_Y\}$ $\delta'(Y_Y, D) = \{Y_Y\}$ $\delta'(\mathcal{S}_Y, b) = \{X_Y, S_Y, B_Y\}$ $\delta'(\mathcal{S}_Y, a) = \{Y_Y, A_Y\}$

Krok 3. Na podstawie powyższych automatów konstruowane są następujące gramatyki:

G'_X	G'_A	G'_B
$\mathcal{S}_X \to bX_X \mid b \mid bB_X \mid aY_X \mid aA_X$ $A_X \to YC_X$ $B_X \to XD_X$ $X_X \to CX_X \mid C$ $Y_X \to DY_X$	$\mathcal{S}_A \to bX_A \mid bB_A \mid aY_A \mid aA_A \mid a$ $A_A \to YC_A$ $B_A \to XD_A$ $X_A \to CX_A$ $Y_A \to DY_A$	$\mathcal{S}_B \to bX_B \mid b \mid bB_B \mid aY_B \mid aA_B$ $A_B \to YC_B$ $B_B \to XD_B$ $X_B \to CX_B$ $Y_B \to DY_B$
Po usunięciu symboli nieużytecznych:		
$\mathcal{S}_X \to bX_X \mid b$ $X_X \to CX_X \mid C$	$\mathcal{S}_A \to a$	$\mathcal{S}_B \to b$
$L(G'_X) = L_X = \{\, bC^k \mid k \geq 0 \,\}$	$L(G'_A) = L_A = \{a\}$	$L(G'_B) = L_B = \{b\}$

G'_C	G'_D	G'_Y
$\mathcal{S}_C \to bX_C \mid bB_C \mid aY_C \mid aA_C$ $A_C \to Y$ $B_C \to XD_C$ $X_C \to CX_C$ $Y_C \to DY_C$	$\mathcal{S}_D \to bX_D \mid bB_D \mid aY_D \mid aA_D$ $A_D \to YC_D$ $B_D \to X$ $X_D \to CX_D$ $Y_D \to DY_D$	$\mathcal{S}_Y \to bX_Y \mid bB_Y \mid aY_Y \mid a \mid aA_Y$ $A_Y \to YC_Y$ $B_Y \to XD_Y$ $X_Y \to CX_Y$ $Y_Y \to DY_Y \mid D$
Po usunięciu symboli nieużytecznych:		
$\mathcal{S}_C \to aA_C$ $A_C \to Y$	$\mathcal{S}_D \to bB_D$ $B_D \to X$	$\mathcal{S}_Y \to aY_Y \mid a$ $Y_Y \to DY_Y \mid D$
$L(G'_C) = L_C = \{aY\}$	$L(G'_D) = L_D = \{bX\}$	$L(G'_Y) = L_Y = \{\, aD^k \mid k \geq 0 \,\}$

Krok 4.

G_X :
$\mathcal{S}_X \to bX_X \mid b$
$X_X \to aA_CX_X \mid aA_C$

G_A :
$\mathcal{S}_A \to a$

G_B :
$\mathcal{S}_B \to b$

G_C :
$\mathcal{S}_C \to aA_C$
$A_C \to aY_Y \mid a$

G_D :
$\mathcal{S}_D \to bB_D$
$B_D \to bX_X \mid b$

G_Y :
$\mathcal{S}_Y \to aY_Y \mid a$
$Y_Y \to bB_DY_Y \mid bB_D$

Krok 5.

G' :
$S \to b \mid bX_XC \mid bC$
$X \to bX_XC \mid bC \mid b$
$C \to aY$
$Y \to aY_YD \mid aD \mid a$
$D \to bX$
$X_X \to aA_CX_X \mid aA_C$
$A_C \to aY_Y \mid a$
$B_D \to bX_X \mid b$
$Y_Y \to bB_DY_Y \mid bB_D$
$A \to a$
$B \to b$

Ostatecznie, po usunięciu symboli nieosiągalnych, gramatyka G' ma postać:

G' :
$S \to b \mid bX_XC \mid bC$
$X \to bX_XC \mid bC \mid b$
$C \to aY$
$Y \to aY_YD \mid aD \mid a$
$D \to bX$
$X_X \to aA_CX_X \mid aA_C$
$A_C \to aY_Y \mid a$
$B_D \to bX_X \mid b$
$Y_Y \to bB_DY_Y \mid bB_D$

6 Podsumowanie

Metody Rosenkrantz'a i Autebert'a są stosunkowo skomplikowane dla gramatyk z większą liczbą nieterminali. Metoda Autebert'a wydaje się łatwiejsza i mniej podatna na błędy od metody Rosenkrantz'a. Jest to spowodowane głównie przez wzorce pokazane w najbardziej skomplikowanej części metody t.j. domknięcie względem lewostronnego oraz prawostronnego ilorazu. Reszta metody jest prosta. Podobne wzorce nie występują w metodzie Rosenkrantz'a.

Możliwości zaprezentowanych algorytmów podsumowuje poniższa tabela. Kwadratowa postać normalna Greibach (2-GNF) jest niejako graniczną postacią dla gramatyk bezkontekstowych. Uzyskanie bardziej standardowej postaci 1-GNF oznaczałoby że gramatyka jest regularna.

	GNF	2-GNF	podwójna GNF	podwójna 2-GNF
Paull	$\surd$	-	-	-
Rosenkrantz	$\surd$	-	$\surd$	-
Autebert	$\surd$	-	$\surd$	$\surd$
Blum-Koch	$\surd$	$\surd$	-	-

Literatura

[1] Blum N., Koch R., *Greibach Normal Form Transformation*, Informatik IV, Universität Bonn.

[2] Autebert J.-M., Berstel J., Boasson L., *Context-Free Languages and Pushdown Automata*, Université Denis Diderot, Paris. Université Pierre et Marie Curie, Paris.

[3] Rosenkrantz D. J., *Matrix Equations and Normal Forms for Context-Free Grammars*, Columbia University, New York.