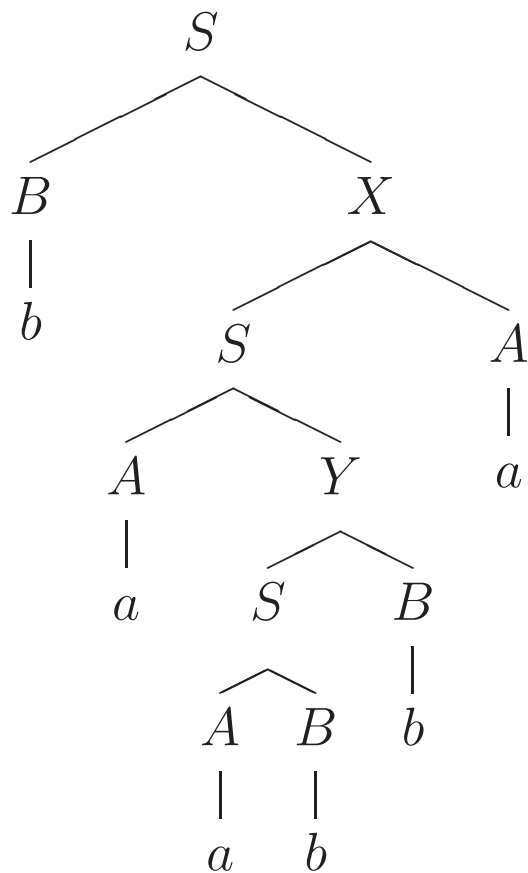


復習：導出木

a と b を同じ数含む長さ 1 以上の記号列を生成する文法のチョムスキー標準形の生成規則

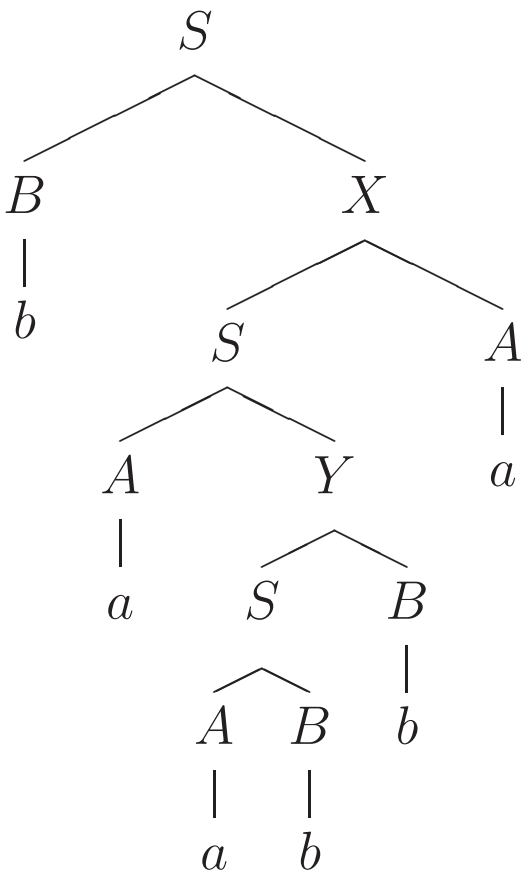
$$P = \{S \rightarrow AB|BA|AY|BX|SS, \\ X \rightarrow SA, Y \rightarrow SB, A \rightarrow a, B \rightarrow b\}$$

$baabba$ の導出木



上記のような表記法を導出木と呼ぶ。

文脈自由文法のポンプの補題の例



上の例で $u = b$, $v = a$, $w = ab$, $x = b$, $y = a$ と置くと、まんなかにある S を書き換える操作を繰り返して、

$$uv^iwx^iy = ba^iabb^ia$$

も導出できることがわかる。このことを一般的に書くと…

文脈自由文法のポンプの補題

L が文脈自由言語ならば、ある n が存在し、 $|z| \geq n$ である L の記号列 z について以下の条件を満たす記号列 u, v, w, x, y が存在する

1. $z = uvwxy$
2. $|vx| \geq 1$
3. $|vwx| \leq n$
4. すべての $i \geq 0$ について $uv^iwx^iy \in L$

証明：

$G: L \setminus \{\epsilon\}$ を生成するチョムスキー標準形

m : G の非終端記号の数

$$n = 2^m$$

導出木の高さを、開始記号から終端記号に至る枝の数の最大値とする。高さが h の導出木を持つ記号列の長さは 2^{h-1} 以下である。従って $|z| \geq n$ なら z の導出木の高さは $m+1$ 以上である。導出木の開始記号から終端記号に至る経路で最も長いものには $m+1$ 個以上の非終端記号がある。2回以上現れる非終端記号のうち終端記号に最も近いものを A として、

$$S \xRightarrow{*} uAy,$$

$$A \xRightarrow{*} vAx,$$

$$A \xRightarrow{*} w,$$

と u, v, w, x, y を定義すると (教科書 85 ページの図 5.3 を参照)、

1. $z = uvwxy$
2. 導出木の上にある A は 2 つの子を持つので $|vx| \geq 1$
3. (終端記号近い非終端記号を選んだので) 導出木の上にある A から終端記号まで枝の数は $m + 1$ 以下である。従って $|vwx| \leq 2^m = n$
4. $A \xRightarrow{*} vAx$ の導出を i 回繰り返すと uv^iwx^iy を導出できる

ポンプの補題の使用例

$L = \{a^m b^m c^m \mid m \geq 1\}$ は文脈自由文法ではない。

ある n が存在して

1. $a^n b^n c^n = uvwxy$,
2. $|vx| \geq 1$,
3. $|vwx| \leq n$,
4. すべての i について $uv^i wx^i y \in L$

とできたとして矛盾を導く。条件 3 より vx が a, b, c すべてを含むことはあり得ない。従って uv^2wx^2y は a と b と c の数が等しくないので L に含まれず矛盾。

演習問題 49 $\{a^m b^m c^i \mid m \geq 1, i \leq m\}$ が文脈自由言語ではないことを証明せよ。

演習問題 50 $\{ww \mid w \in \{a, b\}^+\}$ が文脈自由言語ではないことを証明せよ。

問題 51 今日の授業でわかりにくい所や要望を書いて下さい