3.5 文脈自由文法の性質

- この節の内容
  - 定理3.8　　ε生成規則消去定理
  - 定理3.9　　鎖生成規則の消去定理
  - 定理3.10　Chomsky標準形への変形
  - 定理3.11　挿入定理

定理3.8 ε 生成規則消去定理

定理3.8

- 文脈自由文法 $G = (N, \Sigma, P, S)$ に対して、次の (1)ー(3) を満たす文脈自由文法 $G' = (N', \Sigma, P', S')$ を構成できる。

(1) $L(G') = L(G)$

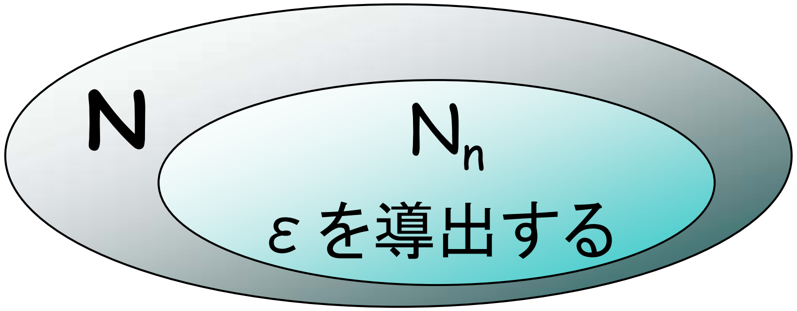
(2) $\varepsilon \in L(G)$ のときのみ G' は ε 生成規則を持ち、
それは $S' \to \varepsilon$ に限る。

(3) G' の開始記号 S' は P' のどの生成規則の
右辺にも現れない.

証明の手引き

- G に対して、条件に合うような G' を定義する。
  - その前準備として、元の G の非終端記号 N のうち ε を生成するものの集合 N_n を作る。
- こうして作った G' が (2) と (3) を満たすのは明らか。

- 残る(1) $L(G') = L(G)$ を証明する。
  - $L(G') \subseteq L(G)$ を証明
  - $L(G') \supseteq L(G)$ を証明



証明 (1)

証明

- $N' = N \cup \{S'\}$
- P' は次のとおり　このP'の作り方がε生成規則を省く方法になる

(i)　$\varepsilon \in L(G)$ ならば $S' \to \varepsilon$ は生成規則

(ii)　$S' \to S$ は生成規則

(iii)　$A \to \alpha_1 A_1 \alpha_2 \cdots \alpha_k A_k \alpha_{k+1} \in P$,

$A_i \in N \cup \Sigma\ (1 \le i \le k)$, $\underline{\alpha_k \in N_n{}^*}$

ならば、

$A \to A_1 \cdots A_k$ は生成規則

注: α_k はεを導出する

証明 (2)

- $L(G') \subseteq L(G)$
  - (i) から、$\varepsilon \in L(G) \Leftrightarrow \varepsilon \in L(G')$
    ここで、$S' \overset{*}{\Rightarrow}_{G'} w \quad (w \in \Sigma^+)$ ならば、
    $S' \overset{*}{\Rightarrow}_{G'} S \overset{*}{\Rightarrow}_{G} w$ となる導出がある。
- $L(G') \supseteq L(G)$
  - (b) $A \in N \cup \Sigma$ のとき, $A \Rightarrow_G w, w \in \Sigma^+$ ならば
    $A \Rightarrow_{G'} w$ である。
  - (b) を導出の長さの帰納法によって証明。
  - (b) より, $S \Rightarrow_G w, w \in \Sigma^+$ ならば
    $S \Rightarrow_{G'} w$ となるので、
    $S' \Rightarrow_{G'} w$ となる。

例3.17

- $G = (N, \Sigma, P, S)$ を次の生成規則をもつ文脈自由文法とする。

$$S \to AB$$
$$A \to ABAC \mid B \mid a$$
$$B \to C \mid b \mid \varepsilon$$
$$C \to B \mid c \mid \varepsilon$$

ε 生成規則をなくした G' をつくってみよう

定理3.9 鎖生成規則の消去定理

定理3.9

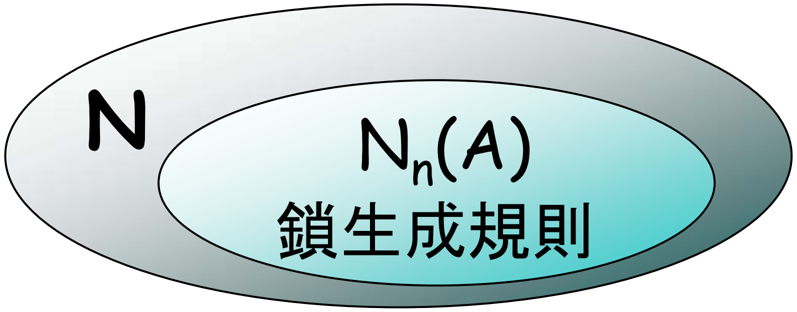
- 文脈自由文法 $G = (N, \Sigma, P, S)$ に対して、
  次の (1)–(3) を満たす文脈自由文法
  $G' = (N, \Sigma, P', S)$ を構成できる。
  (1) $\varepsilon \in L(G)$ のときのみ $S \to \varepsilon$
  (2) $A \to \alpha$ $\quad (\,|\alpha| \geq 2,\ \alpha \in ((N \cup \Sigma - \{S\})^*)$
  (3) $A \to a$ $\quad (\, a \in \Sigma\,)$
  (4) $L(G') = L(G)$

注 : S を右辺に含まないということ

証明の手引き

- G に対して、条件に合うような G' を定義する。
  - その前準備として、もとの G の非終端記号 N のうち、鎖生成規則であるような集合 $N_n(A)$ を作る。
- こうして作った G' が (1)—(3) を満たすのは明らか。

- 残る(4) $L(G') = L(G)$ を証明する。
  - $L(G') \subseteq L(G)$ を証明
  - $L(G') \supseteq L(G)$ を証明



証明 (1)

証明

- N, S は同じ
- P′ は次のとおり、
  $P' = \{A \to \alpha \mid$
  $\qquad {}^{\exists}B \in N_n(A)\ [B \to \alpha \in P$ かつ $\alpha \notin N]\}$
- $N_n(A) = \{B \in N \mid A \overset{*}{\Rightarrow}_G B\} \quad (n = |N|)$

証明 (2)

- $L(G') \subseteq L(G)$
  - $A \to \alpha \in P'$ ならば、ある $B \in N_n(A)$ が存在して、$A \overset{*}{\Rightarrow}_G B \overset{*}{\Rightarrow}_G \alpha$ となる。よって $A \overset{*}{\Rightarrow}_G \alpha$
  - ということは、$A \overset{*}{\Rightarrow}_{G'} w$ ならば必ず $A \overset{*}{\Rightarrow}_G w$ となるので、$L(G') \subseteq L(G)$。
- $L(G') \supseteq L(G)$
  - $A \in N, w \in \Sigma^+$ に対して、$A \overset{+}{\Rightarrow}_G w$ ならば $A \overset{*}{\Rightarrow}_{G'} w$
  - これを導出の長さの帰納法によって証明（略）。
  - これより、$S \overset{+}{\Rightarrow}_G w, w \in \Sigma^+$ ならば $S \overset{*}{\Rightarrow}_{G'} w$ となる。よって $L(G') \supseteq L(G)$。

例3.18

- $G = (N, \Sigma, P, S)$ を次の生成規則をもつ文脈自由文法とする。

$$S \to A$$
$$A \to AB \mid a$$
$$B \to C \mid b$$
$$C \to A \mid c$$

鎖生成規則をなくした G' をつくってみよう

Chomsky標準形

定義

- $G = (N, \Sigma, P, S)$ はその生成規則の形が
  $A \to BC \qquad (A, B, C \in N)$
  または,
  $A \to a \qquad (A \in N, a \in \Sigma)$
  のとき Chomsky標準形であるという。

- 注：Chomsky標準形の文脈自由文法によって生成される言語は，空語 ε を含まない。

定理3.10 Chomsky標準形への変形

定理3.10

- 文脈自由文法 $G = (N, \Sigma, P, S)$ に対して、$L(G) - \{\varepsilon\}$ を生成する Chomsky 標準形の文脈自由文法 $G' = (N', \Sigma, P', S')$ を構成できる。

証明

証明

- 定理3.9により、G の $S \to \varepsilon$ 以外の生成規則の形は

  $A \to \alpha \quad (|\alpha| \geq 2,\ \alpha \in ((N \cup \Sigma - \{S\})^*)$

  $A \to a \quad (a \in \Sigma)$

  であるとしてよい。

  このとき G の生成規則を次のように置き換える。

$$A \to \underbrace{X_1}_{Y_1}\ \underbrace{X_2\ X_3\ X_4\ X_5\ \cdots\ X_k}_{Z_1} \qquad (X_i \in (N \cup \Sigma) - \{S\},\ k \geq 2)$$

$$Z_1 \to Y_2\ Z_2$$

$$Z_2 \to Y_3\ Z_3$$

$$\vdots$$

$$Z_{k-2} \to Y_{k-1} Y_k$$

$X_i \to a$ ならば,

$Y_i \to a$

例3.19

- $G = (N, \Sigma, P, S)$ を次の生成規則をもつ文脈自由文法とする。

$$S \rightarrow AaAA$$

$$A \rightarrow a$$

等価なChomsky標準形の G' をつくってみよう

定理3.11 挿入定理

定理3.11

- $G = (N, \Sigma, P, S)$ を文脈自由文法とする。
  このとき整数 $p(G)$ が存在して、$|z| > p(G)$ なる任意の$z \in L(G)$ に対して、z の分解

  $$z = u\,v\,w\,x\,y$$

  で次の (1)–(3) を満たすものが存在する。

  (1) すべての $q \geq 0$ に対して、

  $$u\,v^q\,w\,x^q\,y \in L(G)$$ である。

  (2) $v \neq \varepsilon$ または $x \neq \varepsilon$ である。

  (3) $|v\,w\,x| \leq p(G)$ である。

証明の手引き

- 適当な $p(G)$ をとる。
- $|z| > p(G)$ を満たす $z \in L(G)$ について、z の導出に対応する構文木を解析する。
- 以上の $p(G)$、z について、(1)(2)(3) がそれぞれ成り立つことを確かめる。

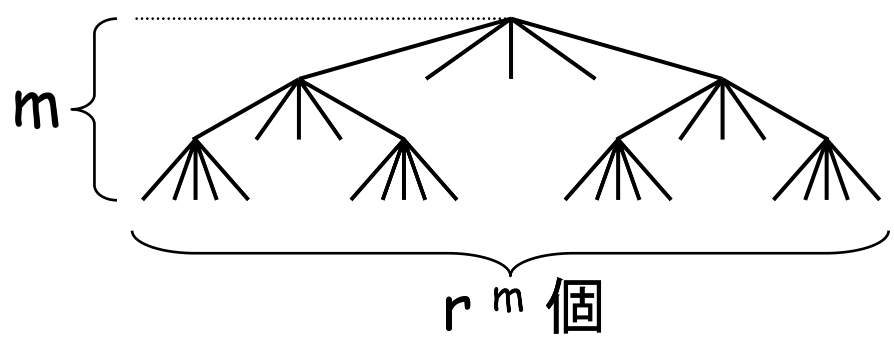
証明（p(G)）

- 構文木の枝分かれの最大数は、

  $$r = \max\{\, |\alpha| \mid A \to \alpha \in P \,\}$$

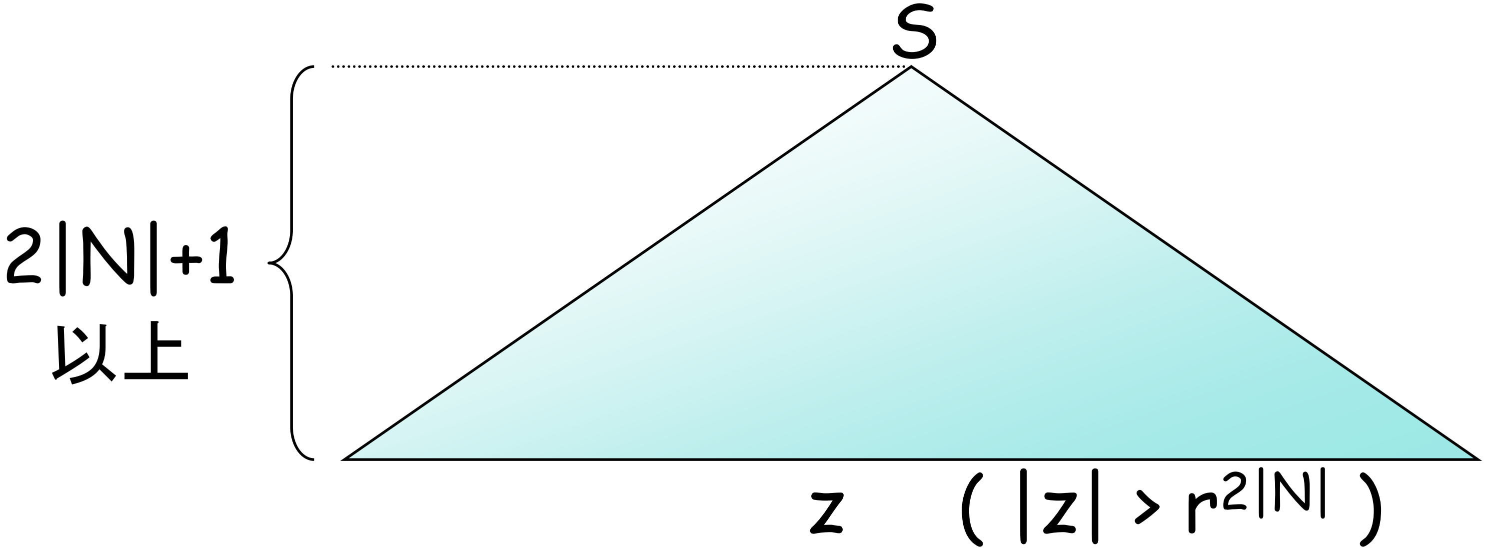
- $m \geq 1$ に対して、高さ m の構文木は
  高々 r^m 個の葉しかもたないことに注意。

- $p(G) = r^{2|N|}$ とおく。

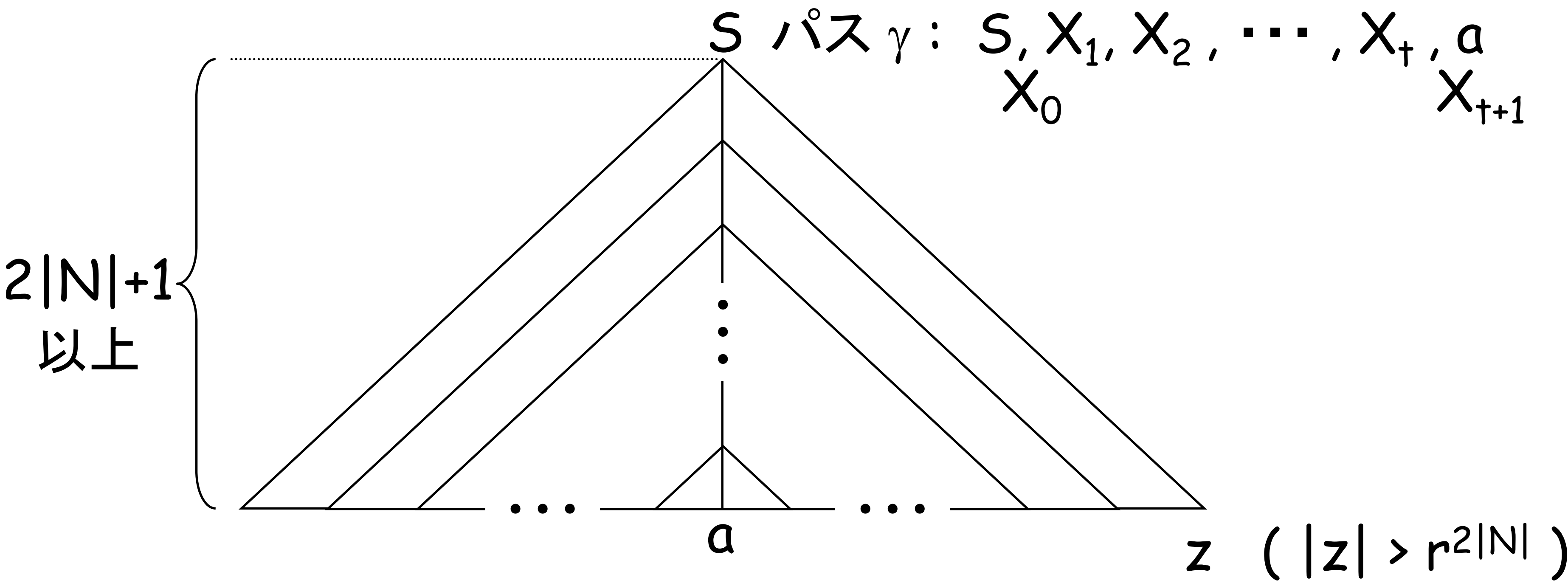


証明（wを導出する構文木）

- $p(G) = r^{2|N|}$
- $z \in L(G)$ は $|z| > p(G)$ を満たしているとする。
- $S \overset{*}{\Rightarrow} z$ を, z を生成する最短の導出と仮定し、この導出に対応する構文木を T とする。



証明（根から葉への最長パスγ）



X_1 を根とする T の部分木を T_i ($0 \leq i \leq t + 1$)
T_i に属する葉の個数を h_i とすると

$$h_0 \geq h_1 \geq \cdots \geq h_{t+1} = 1$$

証明（抽き出し法）

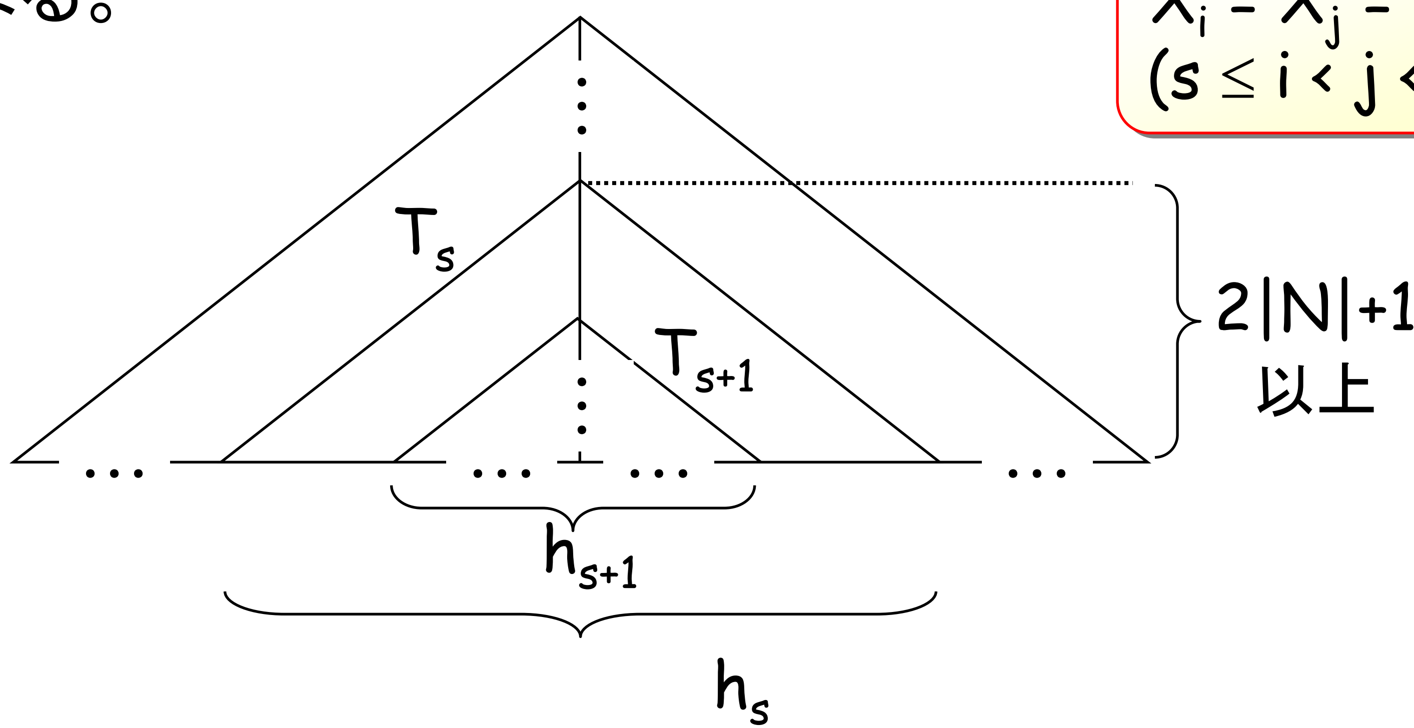
$$h_0 \geq h_1 \geq \cdots \geq h_{t+1} = 1$$

$h_0 > r^{2|N|} \geq 1$ だから、ある s が存在して

$$h_s > r^{2|N|} \geq h_{s+1}$$

となる。

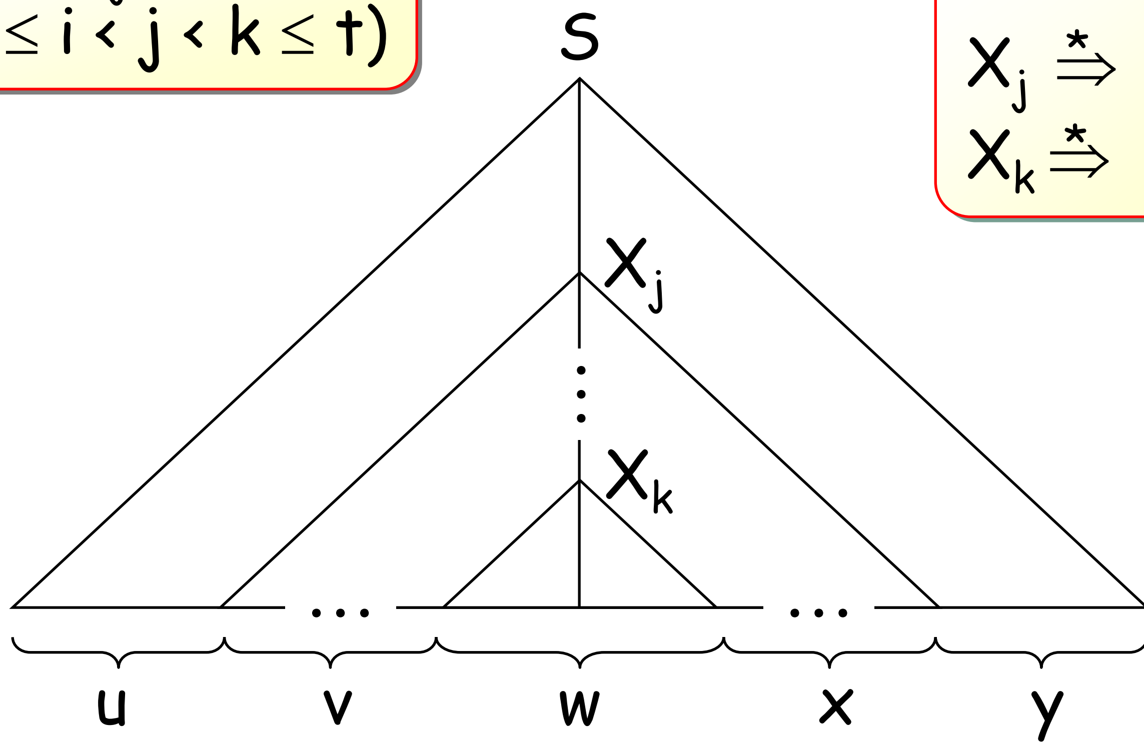
$X_i = X_j = X_k = A$
$(s \leq i < j < k \leq t)$



証明（(3) $|v\,w\,x| \leq p(G)$ ）

$X_i = X_j = X_k = A$
$(s \leq i < j < k \leq t)$

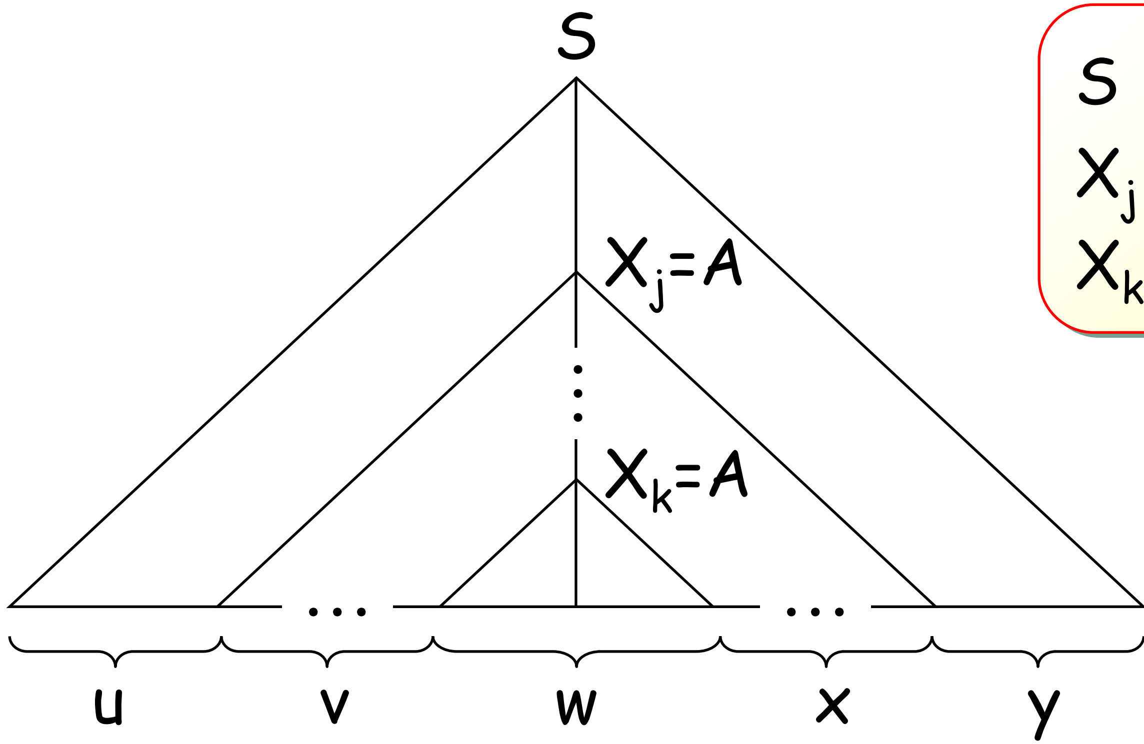
$S \overset{*}{\Rightarrow} u\,X_j\,y$
$X_j \overset{*}{\Rightarrow} v\,X_k\,x$
$X_k \overset{*}{\Rightarrow} w$



$s \leq i < j$ だから

$$|v\,w\,x| \leq h_j \leq h_{s+1} \leq r^{2|N|}$$

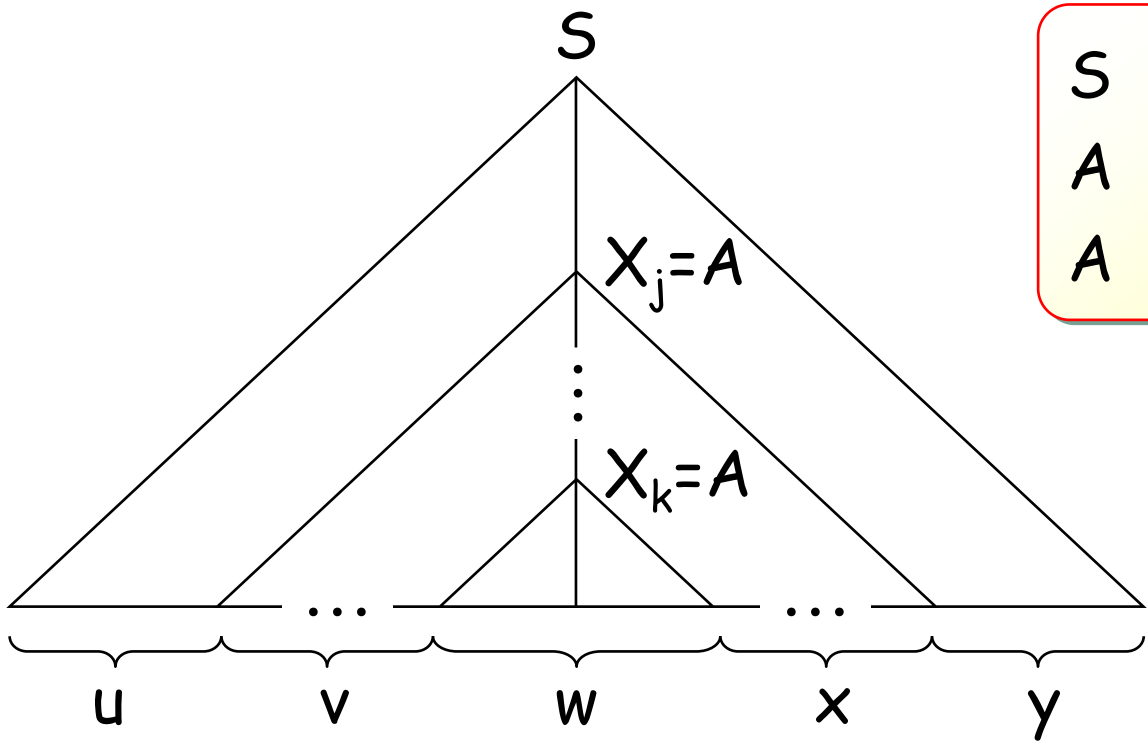
証明((2) $v \neq \varepsilon$ または $x \neq \varepsilon$)



$$S \overset{*}{\Rightarrow} u X_j y$$
$$X_j \overset{*}{\Rightarrow} v X_k x$$
$$X_k \overset{*}{\Rightarrow} w$$

$v = x = \varepsilon$ とすると、$A = X_j = X_k$ であるので、導出 $X_j \overset{*}{\Rightarrow} X_k$ を短絡させて、より短い導出で z を導出できる。
これは、$S \overset{*}{\Rightarrow} z$ が最短の導出であることに矛盾する。

証明((1)$^\forall q \geq 0$ に対して, $u\, v^q\, w\, x^q\, y \in L(G)$)



$S \stackrel{*}{\Rightarrow} u\,A\,y$

$A \stackrel{*}{\Rightarrow} v\,A\,x$

$A \stackrel{*}{\Rightarrow} w$

$^\forall q \geq 0$ について、

$$S \stackrel{*}{\Rightarrow} uAy \stackrel{*}{\Rightarrow} uvAxy \stackrel{*}{\Rightarrow} \cdots \stackrel{*}{\Rightarrow} uv^qAx^qy \stackrel{*}{\Rightarrow} uv^qwx^qy$$

となるので、$u\, v^q\, w\, x^q\, y \in L(G)$

例3.20

- $L = \{a^n b^n c^n \mid n \geq 0\}$ が文脈自由言語でないことを挿入定理を用いて証明せよ。