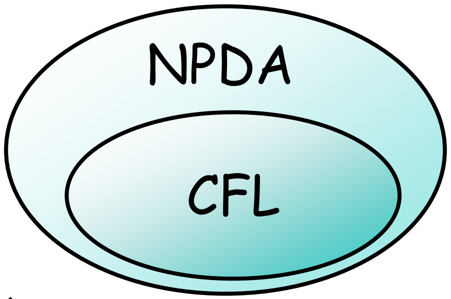
3.3 文脈自由文法とPDA

定理3.4

- CFL = NPDA
  - 文脈自由言語のクラスと非決定性プッシュダウンオートマトンによって受理される言語のクラスは一致する。
- 2つの補題に分けてこの定理を証明する。
  - 補題3.5 CFL⊆NPDA
  - 補題3.6 NPDA⊆CFL

補題3.5 CFL⊆NPDA

- この補題の意味
  - 任意の言語 $L \in CFL$ について、
    $L \in NPDA$ （L を受理するnpdaがある）



- 証明のアイデア
  - $L \in CFL$とし、$L \subseteq \Sigma^*$ を生成する文脈自由文法を $G = (N, \Sigma, P, S)$ とする
  - この G に対応する npda M をつくり、$L(G) = N(M)$ となることを示せばよい

補題3.5の証明 (1)

- 言語 L を最終状態と空ストアで受理する npda $M = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ を次のように定義
  - $K = \{q_0\}$
  - $\Gamma = N \cup \Sigma$
  - q_0 が初期状態
  - $S \in N$ が初期記号 Z_0
  - $F = \{q_0\}$
  - 遷移関係δ は次のように与えられる。

    (i) $\delta(q_0, \varepsilon, A) = \{(q_0, \alpha) \mid A \rightarrow \alpha \in P\}$

    (ii) $\delta(q_0, a, a) = \{(q_0, \varepsilon)\}$ (ただし $a \in \Sigma$)

注:これは、文脈自由文法から等価なPDAを作る手法になる!

補題3.5の証明 (2)

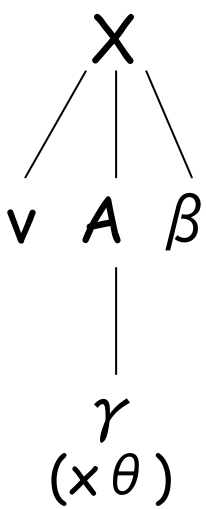
- $L(G) = N(M)$ を証明するぞ！
  - まずは $X \in N$, $u \in \Sigma^*$, $\alpha \in N(N \cup \Sigma)^* \cup \{\varepsilon\}$ とするとき次の (a) (b) が同値であることを示す。
    - (a) 最左導出 $X \overset{*}{\Rightarrow} u\alpha$ がある。
    - (b) $(q_0, u, X) \vdash^* (q_0, \varepsilon, \alpha)$ となる。

      入力から u を消費してストアのXを α に変える
  - (a)⇒(b)（導出の長さについての帰納法）
    - 長さが 0 のときは $u = \varepsilon$, $\alpha = X$ であるので自明。
    - 最左導出の長さが n のとき成立すると仮定する。
      このとき長さ n+1の最左導出

      $$X \overset{n+1}{\Rightarrow} u\alpha$$

      をとる。ただし $u \in \Sigma^*$, $\alpha \in (N \cup \Sigma)^* \cup \{\varepsilon\}$ とする。

補題3.5の証明 (3)



この最左導出は、

$X \stackrel{n}{\Rightarrow} v\underline{A\beta} \Rightarrow v\gamma\beta$ (ただし $v \in \Sigma^*$, $A \rightarrow \gamma \in P$)

のように、長さ n の最左導出と、長さ 1 の最左導出に分解できる。

ここで $u = vx$, $x \in \Sigma^*$, $\gamma\beta = x\alpha$ と書ける。帰納法の仮定より

$$(q_0, v, X) \vdash^* (q_0, \varepsilon, A\beta)$$

となる。したがって

$$(q_0, vx, X) \vdash^* (q_0, x, A\beta)$$
$$\vdash (q_0, x, x\alpha)$$ ((i)の遷移と $\gamma\beta = x\alpha$ による)
$$\vdash^* (q_0, \varepsilon, \alpha)$$ ((ii)の遷移による)

このようにして $(q_0, u, X) \vdash (q_0, \varepsilon, \alpha)$となる。

よって(a)⇒(b)が成り立つ。

補題3.5の証明 (4)

- (b)⇒(a)の証明
  - 計算のステップ数についての帰納法で証明する。
  - ステップ数が 0 の場合は自明。
  - ステップ数が n 以下のとき成立すると仮定する。長さ n+1 の計算
    $$(q_0, u_0, a_0) \vdash \cdots \vdash (q_0, u_n, a_n) \vdash (q_0, u_{n+1}, a_{n+1})$$
    をとる。ただし $u_0 = u$, $u_{n+1} = \varepsilon$, $a_0 = X$, $a_{n+1} = a$ とする。
  - n+1 番目のステップで (i) の遷移（ε 動作）が使われている場合：
    上の計算は
    $$(q_0, u, X) \vdash^n (q_0, \varepsilon, A\beta) \vdash (q_0, \varepsilon, \gamma\beta)$$
    と分解できる。ただし $A \to \gamma \in P$, $\gamma\beta = \alpha$ である。
    帰納法の仮定により、最左導出
    $$X \overset{*}{\Rightarrow} uA\beta$$
    が存在する。$A \to \gamma \in P$ であり、また $u \in \Sigma^*$ だから、
    $$X \overset{*}{\Rightarrow} uA\beta \Rightarrow u\gamma\beta$$
    は最左導出である。
    よって $X \overset{*}{\Rightarrow} u\alpha$ なる最左導出が存在する。

補題3.5の証明 (5)

- n+1 番目のステップで (ii) の遷移が使われている場合：
  X は非終端記号なので、第1ステップ目では (ii) の遷移を適用できない。したがって、ある時点 m($2 \leqq m \leqq n+1$) が存在して、m－1ステップ目では (i)の遷移が使われ、すべての t($m \leqq t \leqq n+1$)に対して t ステップ目では (ii) の遷移が適応されている。
- すると、計算 $(q_0, u, X) \vdash^{n+1} (q_0, \varepsilon, \alpha)$ は (u = vx とおくと)

$$
\begin{aligned}
(q_0, vx, X) &\vdash^{m-2} (q_0, x, A\beta) \\
&\vdash (q_0, x, \gamma\beta) \quad \text{（m-1回目の遷移は(i)の遷移）} \\
&\vdash^{n-m+2} (q_0, \varepsilon, \alpha) \quad \text{（残りはすべて(ii)の遷移）}
\end{aligned}
$$

  と分解できる。また $(q_0, vx, X) \vdash^{m-2} (q_0, x, A\beta)$ ならば $(q_0, v, X) \vdash^{m-2} (q_0, \varepsilon, A\beta)$ であるので、帰納法の仮定より最左導出

$$X \overset{*}{\Rightarrow} vA\beta$$

  が存在する。すると　（m-1回目の遷移に対応）

$$X \overset{*}{\Rightarrow} vA\beta \Rightarrow v\gamma\beta$$

  は最左導出である。
  $\gamma\beta = x\alpha$, $vx = u$ であるので $X \overset{*}{\Rightarrow} u\alpha$ があることになる。

補題3.5の証明 (6)

- 以上により (a) と (b) が同値であることがわかった。
  (a) 最左導出 $X \overset{*}{\Rightarrow} u\alpha$ がある。
  (b) $(q_0, u, X) \vdash^* (q_0, \varepsilon, \alpha)$ となる。
- 補題3.1より、G の導出は最左導出としておいてよいので、
  $w \in \Sigma^*$ に対して
  $$S \overset{*}{\Rightarrow} w \quad \Leftrightarrow \quad (q_0, w, S) \vdash^* (q_0, \varepsilon, \varepsilon)$$
  となる。したがって L(G) = N(M) となる。
  命題3.2より $L \in$ NPDA となる。　　　　　　【証明終】

補題3.1
$G = (N, \Sigma, P, S)$ を文脈自由文法とする。導出 $u \overset{*}{\Rightarrow} v$ があるならば、u から v への最左導出がある。

例3.13

- 文脈自由文法 $G = (N, \Sigma, P, S)$ を
  - $N = \{S\}$
  - $\Sigma = \{a, b\}$
  - $P = \{S \rightarrow aSb,\ S \rightarrow ab\}$

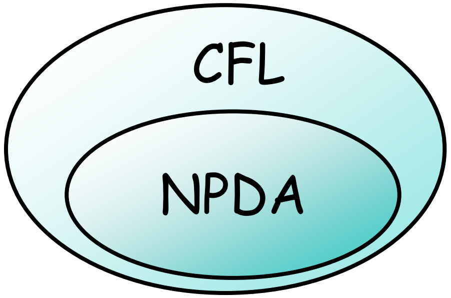
とする。このとき補題3.5で構成されるnpdaの遷移関数 δ は次のようになる

現在の状態　入力ヘッダが見ている記号　ストアのヘッダが見ている記号

q	x	Z	$\delta(q, x, Z)$
q_0	ε	S	$\{(q_0, aSb), (q_0, ab)\}$
q_0	a	a	$\{(q_0, \varepsilon)\}$
q_0	b	b	$\{(q_0, \varepsilon)\}$

補題3.6 NPDA⊆CFL

- この補題の意味
  - 任意の言語 L∈NPDA について、 L∈CFL （L を導出するCFLがある）



- 証明のアイデア
  - L∈NPDAとし、この L を最終状態と空ストアで受理するnpdaを $M=(K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ とする
  - この M に対応する文脈自由文法 G をつくり、N(M) = L(G) となることを示せばよい

補題3.6の証明 (1)

- L を生成する文脈自由文法 $G=(N, \Sigma, P, S)$ を次のように定義
  - $N = (K \times \Gamma \times K) \cup \{S\}$
  - P は次の生成規則よりなる

(i) 各 $p \in F$ に対して $S \to (q_0, Z_0, p)$ は生成規則である。

(ii) $(p, Z_1 \cdots Z_k) \in \delta(q, a, Z)$ $(k \geqq 1, a \in S \cup \{\varepsilon\})$のとき
任意の$q_1, q_2, \ldots, q_k \in K$ に対して
$(q, Z, q_k) \to a(p, Z_1, q_1)(q_1, Z_2, q_2) \cdots (q_{k-1}, Z_k, q_k)$
は生成規則である

(iii) $(p, \varepsilon) \in \delta(q, a, Z)$ $(a \in S \cup \{\varepsilon\})$ のとき
$(q, Z, p) \to a$ は生成規則

注:これは、PDAから等価な文脈自由文法を作る手法になる！

補題3.6の証明 (2)

- L(G) = N(M) を証明するぞ！
  - その前に、次の (a) (b) が同値であることを示す。
    (a) $(q, Z, p) \overset{*}{\Rightarrow} x$
    (b) $(q, x, Z) \vdash^{*} (p, \varepsilon, \varepsilon)$
  - (a)⇒(b)の証明（導出の長さについての帰納法で証明）
    - 長さが 1 のときは、生成規則 (iii) により
      $x \in S \cup \{\varepsilon\}$ で $(p, \varepsilon) \in d(q, x, Z)$
      となっている。したがって $(q, x, Z) \vdash (p, \varepsilon, \varepsilon)$ となる。
      長さ n 以下の導出に対して成立することを仮定する。
    - 長さ n+1 ≧ 2 の導出
      $(q, Z, p) \Rightarrow x$
      をとる。導出の長さが 2 以上だから一番初めに適用される生成規則は (ii) の形のものである。

補題3.6の証明 (3)

このとき上の導出は

$(q, Z, p) \Rightarrow a(q_1, Z_1, q_2)(q_2, Z_2, q_3)\cdots(q_k, Z_k, p)$
$\quad\quad \overset{*}{\Rightarrow} x$

と書ける。ただし $(q_1, Z_1\cdots Z_k) \in \delta(q, a, Z)$ である。

すると、$x = ax_1\cdots x_k$ $(x_i \in \Sigma^*)$ と書けて、各 i $(1 \leqq i \leqq k)$に対して、長さ n 以下の導出で

$$(q_i, Z_i, q_{i+1}) \overset{*}{\Rightarrow} x_i$$

となる。ただし $q_{k+1} = p$ とする。

帰納法の仮定より

$$(q_i, x_i, Z_i) \vdash (q_{i+1}, \varepsilon, \varepsilon) \quad\quad (1 \leqq i \leqq k)$$

となる。このとき

$$\begin{aligned}(q, x, Z) &\vdash (q_1, x_1\cdots x_k, Z_1\cdots Z_k)\\ &\vdash^* (q_2, x_2\cdots x_k, Z_2\cdots Z_k)\\ &\quad\cdots\\ &\vdash^* (q_k, x_k, Z_k)\\ &\vdash^* (p, \varepsilon, \varepsilon)\end{aligned}$$

となる。

補題3.6の証明 (4)

- (b)⇒(a)の証明（計算のステップ数についての帰納法）
  - (iii) より $(q, Z, p) \to x$ は P の要素となる。
    よってステップ数が 1 のときは成立する。$n+1 \geqq 2$ として
    $$(q, x, Z) \vdash^{n+1} (p, \varepsilon, \varepsilon)$$
    とする。これを最初の1ステップと残りの n ステップに分解する。
    $n \geqq 1$ だから第 1 ステップでは Z がポップされてεに置き換えられることはない。よって
    $$\begin{aligned}(q, x, Z) &\vdash (r, y, Z_1 \cdots Z_k) \\ &\vdash^{n} (p, \varepsilon, \varepsilon)\end{aligned}$$
    となる。ここで $x = ay$, $a \in \Sigma \cup \{\varepsilon\}$ であり、
    $(r, Z_1 \cdots Z_k) \in \delta(q, a, Z)$ である。

補題3.6の証明 (5)

- 各 Z_i はいずれポップされるので、
  分解 $y = y_1 \cdots y_k$ ($y_i \in \Sigma^*$, $1 \leqq i \leqq k$) と状態 $q_1 \cdots q_k$ が存在して、
  各 i ($1 \leqq i \leqq k$) に対して、n 以下のステップ数で
  $$(q_i, y_i, Z_i) \vdash^* (q_{i+1}, \varepsilon, \varepsilon)$$
  となる。ただし $q_1 = r$, $q_{k+1} = p$ とする。
  帰納法の仮定より
  $$(q_i, Z_i, q_{i+1}) \overset{*}{\Rightarrow} y_i \qquad (1 \leqq i \leqq k)$$
  となる。(ii)より
  $$(q, Z, p) \Rightarrow a(q_1, Z_1, q_2) \cdots (q_k, Z_k, p)$$
  だから
  $$(q, Z, p) \Rightarrow a(q_1, Z_1, q_2) \cdots (q_k, Z_k, p) \overset{*}{\Rightarrow} ay_1 \cdots y_k$$
  となる。よって $(q, Z, p) \overset{*}{\Rightarrow} x$ となる。

補題3.6の証明 (6)

- 以上により (a) と (b) が同値であることがわかった。
  (a) $(q, Z, p) \overset{*}{\Rightarrow} x$
  (b) $(q, x, Z) \vdash^{*} (p, \varepsilon, \varepsilon)$

- そこで $x \in \Sigma^*$ に対して $S \overset{*}{\Rightarrow} x$ ならば (i) により、
  ある状態 $p \in F$ が存在して、$S \Rightarrow (q_0, Z_0, p) \overset{*}{\Rightarrow} x$ となる。
  したがって
  $$(q_0, x, Z_0) \vdash^{*} (p, \varepsilon, \varepsilon)$$
  となり、x は M によって受理される。
  逆に x が M によって受理されれば、
  $S \overset{*}{\Rightarrow} x$ となることも同様にわかる。　　【証明終】

例3.14

- $M = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ を言語 $\{a^n b^n \mid n \geqq 1\}$ を最終状態と空ストアで受理する例3.12の npda とする。このとき補題3.6で構成される文脈自由文法 $G=(N, S, P, S)$ は次のようになる。

$$N = \{q_0, q_1\} \times \{A, Z_0\} \times \{q_0, q_1\} \cup \{S\}$$

P の生成規則は次のとおり。

$$\left.\begin{array}{l}
S \to (q_0, Z_0, q_1) \\
(q_0, Z_0, q_0) \to a(q_0, A, q_0) \\
(q_0, Z_0, q_1) \to a(q_0, A, q_1) \\
(q_0, A, q_0) \to a(q_0, A, q_0)(q_0, A, q_0) \\
(q_0, A, q_0) \to a(q_0, A, q_1)(q_1, A, q_0) \\
(q_0, A, q_1) \to a(q_0, A, q_0)(q_0, A, q_1) \\
(q_0, A, q_1) \to a(q_0, A, q_1)(q_1, A, q_1) \\
(q_0, A, q_1) \to b \\
(q_0, A, q_1) \to b
\end{array}\right\}$$

無駄な生成規則が多いよ (・∀・)

3.3節の結論

- 以下の補題が証明された
  - 補題3.5 CFL⊆NPDA
  - 補題3.6 NPDA⊆CFL
- すなわち

定理3.4

CFL = NPDA

- 文脈自由言語のクラスと非決定性プッシュダウンオートマトンによって受理される言語のクラスは一致する。（言語を定義する能力が等しい）