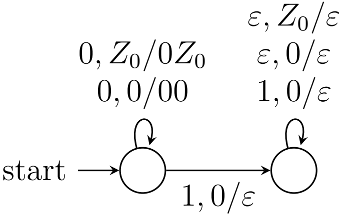
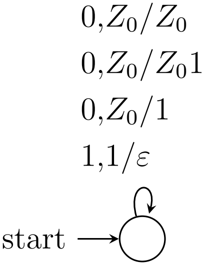
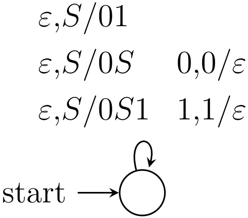
- 下推自动机
- 下推自动机接受的语言
- 下推自动机与文法的等价性
  - 由 CFG 到 PDA
  - 由 PDA 到 CFG
- 确定型下推自动机

由 CFG 到 PDA

例 5. 设计语言 $L = \{0^n 1^m \mid 1 \le m \le n\}$ 的 PDA.







续例 5. 设计语言 $L = \{0^n 1^m \mid 1 \le m \le n\}$ 的 CFG.

CFG G:

$$
\begin{aligned}
S &\to AB \\
A &\to 0A \mid \varepsilon \\
B &\to 0B1 \mid 01
\end{aligned}
$$

字符串 00011 的最左派生:

$$S \underset{\text{lm}}{\Rightarrow} AB \underset{\text{lm}}{\Rightarrow} 0AB \underset{\text{lm}}{\Rightarrow} 0B \underset{\text{lm}}{\Rightarrow} 00B1 \underset{\text{lm}}{\Rightarrow} 00011$$

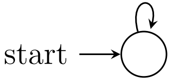
续例 5. 语言 $L=\{0^n1^m \mid 1\le m\le n\}$.
用 PDA 栈顶符号的替换, 模拟文法的最左派生:

PDA		CFG	
PDA 的 ID 转移	PDA 的动作	产生式	最左派生
$(q_0,\ 00011,\ S)$			S
$\vdash (q_0,\ 00011,\ AB)$	$\varepsilon, S/AB$	$S\to AB$	$\underset{\text{lm}}{\Rightarrow} AB$
$\vdash (q_0,\ 00011,\ 0AB)$	$\varepsilon, A/0A$	$A\to 0A$	$\underset{\text{lm}}{\Rightarrow} 0AB$
$\vdash (q_0,\ 0011,\ AB)$	$0, 0/\varepsilon$		
$\vdash (q_0,\ 0011,\ B)$	$\varepsilon, A/\varepsilon$	$A\to\varepsilon$	$\underset{\text{lm}}{\Rightarrow} 0B$
$\vdash (q_0,\ 0011,\ 0B1)$	$\varepsilon, B/0B1$	$B\to 0B1$	$\underset{\text{lm}}{\Rightarrow} 00B1$
$\vdash (q_0,\ 011,\ B1)$	$0, 0/\varepsilon$		
$\vdash (q_0,\ 011,\ 011)$	$\varepsilon, B/01$	$B\to 01$	$\underset{\text{lm}}{\Rightarrow} 00011$
$\vdash (q_0,\ 11,\ 11)$	$0, 0/\varepsilon$		
$\vdash (q_0,\ 1,\ 1)$	$1, 1/\varepsilon$		
$\vdash (q_0,\ \varepsilon,\ \varepsilon)$	$1, 1/\varepsilon$		

续例 5. 语言 $L = \{0^n 1^m \mid 1 \le m \le n\}$.

$$
\begin{aligned}
S &\to AB \\
A &\to 0A \mid \varepsilon \\
B &\to 0B1 \mid 01
\end{aligned}
$$

$$
\begin{array}{lll}
\varepsilon, S/AB & & \\
\varepsilon, A/0A & \varepsilon, A/\varepsilon & 0, 0/\varepsilon \\
\varepsilon, B/0B1 & \varepsilon, B/01 & 1, 1/\varepsilon
\end{array}
$$



定理 27

任何 CFL L, 一定存在 PDA P, 使 $L = \mathbf{N}(P)$.

构造与文法等价的 PDA

如果 CFG $G = (V, T, P', S)$, 构造 PDA

变元和终结符

$$P = (\{q\}, T, V \cup T, \delta, q, S, \varnothing),$$

其中 δ 为:

❶ $\forall A \in V$: 变元

$$\delta(q, \varepsilon, A) = \{(q, \beta) \mid A \to \beta \in P'\}$$

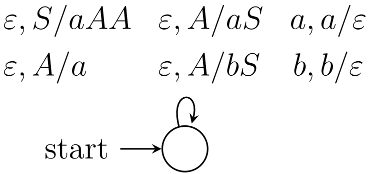
空转移动作

❷ $\forall a \in T$: 终结符

$$\delta(q, a, a) = \{(q, \varepsilon)\}$$

那么 $\mathbf{L}(G) = \mathbf{N}(P)$.

例 6. 为文法 $S \to aAA$, $A \to aS \mid bS \mid a$ 构造 PDA.



证明: 往证

$$S \overset{*}{\Rightarrow} w \Longleftrightarrow (q, w, S) \vdash_{\mathrm{P}}^{*} (q, \varepsilon, \varepsilon).$$

[充分性] 往证

$$S \underset{\mathrm{lm}}{\overset{*}{\Rightarrow}} w \implies (q, w, S) \vdash^{*} (q, \varepsilon, \varepsilon).$$

设 $S \underset{\mathrm{lm}}{\overset{*}{\Rightarrow}} w$ 中第 i 个左句型为 $x_i A_i \alpha_i$, 其中 $x_i \in \Sigma^*, A_i \in V$, $\alpha_i \in (V \cup T)^*$. 并将 S 看作第 0 个左句型 $x_0 A_0 \alpha_0 = S$, 那么

$$x_0 = \varepsilon, A_0 = S, \alpha_0 = \varepsilon.$$

将 w 看作为第 n 个左句型 $x_n A_n \alpha_n = w$, 那么

$$x_n = w, A_n = \varepsilon, \alpha_n = \varepsilon.$$

再对派生步骤 i 归纳, 往证

$$S \underset{\mathrm{lm}}{\overset{i}{\Rightarrow}} x_i A_i \alpha_i \wedge w = x_i y_i \implies (q, w, S) \vdash^{*} (q, y_i, A_i \alpha_i).$$

归纳基础: 当最左派生要 0 步时, 显然成立

$$(q, w, S) \vdash^* (q, y_0, A_0\alpha_0) = (q, w, S).$$

归纳递推: 假设 i 步时上式成立. 当第 $i+1$ 步时,
一定是 $A_i \to \beta$ 应用到 $x_iA_i\alpha_i$

$$S \xRightarrow[\text{lm}]{i} x_iA_i\alpha_i \xRightarrow[\text{lm}]{} x_i\beta\alpha_i = x_{i+1}A_{i+1}\alpha_{i+1}.$$

变元 A_{i+1} 一定在 $\beta\alpha_i$ 中. 设 A_{i+1} 之前的终结符为 x', 则有

$$\beta\alpha_i = x'A_{i+1}\alpha_{i+1}.$$

又因为 $w = x_iy_i = x_ix'y_{i+1} = x_{i+1}y_{i+1}$, 所以有

$$y_i = x'y_{i+1}.$$

那么, 在 PDA 中从 ID $(q, y_i, A_i\alpha_i)$ 模拟最左派生,
用产生式 $A_i \to \beta$ 替换栈顶 A_i 后, 有

$$\begin{aligned}
(q, w, S) &\vdash^* (q, y_i, A_i\alpha_i) && \text{归纳假设} \\
&\vdash (q, y_i, \beta\alpha_i) && A_i \to \beta \\
&= (q, x'y_{i+1}, x'A_{i+1}\alpha_{i+1}) && y_i = x'y_{i+1} \\
&\vdash^* (q, y_{i+1}, A_{i+1}\alpha_{i+1}) && \text{弹出终结符}
\end{aligned}$$

因此 $S \underset{lm}{\overset{n}{\Rightarrow}} w \Longrightarrow (q, w, S) \vdash^* (q, y_n, A_n\alpha_n) = (q, \varepsilon, \varepsilon)$, 即充分性得证.

[必要性] 往证更一般的, 对任何变元 A, 都有:

$$(q, x, A) \vdash^* (q, \varepsilon, \varepsilon) \Longrightarrow A \overset{*}{\Rightarrow} x.$$

可以看作“从输入带中消耗掉 x”与“从栈中弹出 A”两种作用相互抵消.

对 ID 转移 $(q, x, A) \vdash^i (q, \varepsilon, \varepsilon)$ 的次数 i 归纳证明.

归纳基础: 当 $i=1$ 次时, 只能是 $x=\varepsilon$ 且 $A\to\varepsilon$ 为产生式, 所以 $A\overset{*}{\Rightarrow}\varepsilon$.

归纳递推: 假设 $i\le n\ (n\ge 0)$ 时上式成立. 当 $i=n+1$ 时,
因为 A 是变元, 其第 1 步转移一定是

$$(q,x,A)\vdash(q,x,Y_1Y_2\cdots Y_m)$$

且 $A\to Y_1Y_2\cdots Y_m$ 是产生式, 其中 Y_i 是变元或终结符. 而其余的 n 步转移

$$(q,x,Y_1Y_2\cdots Y_m)\overset{*}{\vdash}(q,\varepsilon,\varepsilon)$$

中每个 Y_i 从栈中被完全弹出时, 将消耗掉的那部分 x 记为 x_i, 那么显然有

$$x=x_1x_2\cdots x_m.$$

而每个 Y_i 从栈中被完全弹出时, 都不超过 n 步, 所以由归纳假设,

$$(q,x_i,Y_i)\overset{*}{\vdash}(q,\varepsilon,\varepsilon)\Longrightarrow Y_i\overset{*}{\Rightarrow}x_i.$$

再由 A 的产生式 $A \to Y_1Y_2\cdots Y_m$, 有

$$\begin{aligned} A &\Rightarrow Y_1Y_2\cdots Y_m \\ &\overset{*}{\Rightarrow} x_1Y_2\cdots Y_m \\ &\overset{*}{\Rightarrow} x_1x_2\cdots Y_m \\ &\overset{*}{\Rightarrow} x_1x_2\cdots x_m = x. \end{aligned}$$

因此当 $A = S$, $x = w$ 时,

$$(q, w, S) \overset{*}{\vdash} (q, \varepsilon, \varepsilon) \Longrightarrow S \overset{*}{\Rightarrow} w$$

成立, 即必要性得证.

所以, 任何 CFL 都可由 PDA 识别. $\square$

格里巴赫范式

栈符号表

终结 非终结符

文法开始符号

栈底符号

构造与 GNF 格式文法等价的 PDA

如果 GNF 格式的 CFG $G = (V, T, P', S)$, 那么构造 PDA

$$P = (\{q\},\ T,\ V,\ \delta,\ q,\ S,\ \varnothing),$$

为每个产生式, 定义 δ 为:

$$\delta(q, a, A) = \{(q, \beta) \mid \boxed{A \rightarrow a\beta} \in P'\}.$$

变元 → 终 · 变元V符号串

续例 6. 文法 $S \to aAA$, $A \to aS \mid bS \mid a$ 为 GNF 格式, 构造等价的 PDA.

a: 读掉的终结符

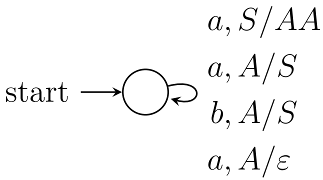
β 可以是 AA

栈顶符号

作为变元

替换的

栈顶



定理 28

如果 PDA P, 有 $L = \mathbf{N}(P)$, 那么 L 是上下文无关语言.

以空栈的方式接收

构造与 PDA 等价的 CFG

如果 PDA $P=(Q,\Sigma,\Gamma,\delta,q_0,Z_0,\varnothing)$, 那么构造 CFG $G=(V,\Sigma,P',S)$, 其中 V 和 P' 为

1. $V=\{[qXp]\mid p,q\in Q,\ X\in\Gamma\}\cup\{S\}$;
2. 对 $\forall p\in Q$, 构造产生式 $S\to[q_0Z_0p]$;
3. 对 $\forall(p,Y_1Y_2\cdots Y_n)\in\delta(q,a,X)$, 构造 $|Q|^n$ 个产生式

$$[qXr_n]\to a[pY_1r_1][r_1Y_2r_2]\cdots[r_{n-1}Y_nr_n]$$

其中 $a\in\Sigma\cup\{\varepsilon\}$, $X,Y_i\in\Gamma$, 而 $r_i\in Q$ 是 n 次 $|Q|$ 种状态的组合; 若 $i=0$, 为 $[qXp]\to a$.

例 7. 将 PDA $P = (\{p, q\}, (0, 1), \{X, Z\}, \delta, q, Z)$ 转为 CFG, 其中 δ 如下:

(1) $\delta(q, 1, Z) = \{(q, XZ)\}$

(2) $\delta(q, 1, X) = \{(q, XX)\}$

(3) $\delta(q, 0, X) = \{(p, X)\}$

(4) $\delta(q, \varepsilon, Z) = \{(q, \varepsilon)\}$

(5) $\delta(p, 1, X) = \{(p, \varepsilon)\}$

(6) $\delta(p, 0, Z) = \{(q, Z)\}$

δ	产生式
(0)	$S \to [qZq]$
	$S \to [qZp]$
(1)	$[qZq] \to 1[qXq][qZq]$
	$[qZq] \to 1[qXp][pZq]$
	$[qZp] \to 1[qXq][qZp]$
	$[qZp] \to 1[qXp][pZp]$
(2)	$[qXq] \to 1[qXq][qXq]$
	$[qXq] \to 1[qXp][pXq]$
	$[qXp] \to 1[qXq][qXp]$
	$[qXp] \to 1[qXp][pXp]$
(3)	$[qXq] \to 0[pXq]$
	$[qXp] \to 0[pXp]$
(4)	$[qZq] \to \varepsilon$
(5)	$[pXp] \to 1$
(6)	$[pZp] \to 0[qZp]$
	$[pZq] \to 0[qZq]$

消除无用符号	重命名 (可选)
$S \to [qZq]$	$S \to A$
$[qZq] \to 1[qXp][pZq]$	$A \to 1BC$
$[qXp] \to 1[qXp][pXp]$	$B \to 1BD$
$[qXp] \to 0[pXp]$	$B \to 0D$
$[qZq] \to \varepsilon$	$A \to \varepsilon$
$[pXp] \to 1$	$D \to 1$
$[pZq] \to 0[qZq]$	$C \to 0A$