1 语言与文法

句子的逆 x^R or x^T.

定义 1.1 文法（Grammar）G 是一个四元组 $G=(V,T,P,S)$ 其中，
V— **变量（Variable）**的非空有穷集。$\forall A\in V$，A 叫做语法变量（syntactic variable），也叫非终极符号（nonterminal）。
T— **终极符（Terminal）**的非空有穷集。$\forall a\in T$，a 叫做终极符。$V\cup T=\emptyset$。
P— **产生式（Production）**的非空有穷集。对于 $a\to b$，a 是**左部**，b 是**右部**。
S— $S\in V$，文法 G 的**开始符号（Start symbol）**。

- 只写产生式，第一个产生式的左部为开始符号
- 对一组有相同左部的产生式
  $\alpha\to\beta_1,\alpha\to\beta_2,\alpha\to\beta_3,\dots$ 可以记为 $\alpha\to\beta_1|\beta_2|\beta_3\dots$ β_1,β_2,β_3 称为**候选式**（Candidate）
- 形如 $\alpha\to\epsilon$ 的产生式叫做空产生式，也叫做 ϵ 产生式
  - 英文大写字母为**语法变量**
  - 英文小写字母为**终结符号**
  - 英文较后的大写字母为**语法变量或者终极符号**
  - 英文较后的大写字母为**终极符号行**
  - 希腊字母表示**语法变量和终极符号组成的行**

定义 1.2 设 $G=(V,T,P,S)$ 是一个文法，如果 $\alpha\to\beta\in P,\gamma,\delta\in(V\cup T)$，则称 $\gamma\alpha\delta$ 在 G 中**直接推导**（Derivation）出 $\gamma\beta\delta$，记作 $\gamma\alpha\delta\underset{G}{\Rightarrow}\gamma\beta\delta$。
于此相对应，$\gamma\beta\delta$ 归约到 $\gamma\alpha\delta$，简称 β 归约为 α。
$\underset{G}{\Rightarrow}$ 是 $(V\cup T)^*$ 上的二元关系。

定义 1.3 对于文法 G：$\overset{n}{\underset{G}{\Longrightarrow}}=\left(\underset{G}{\Rightarrow}\right)^n$ $\overset{*}{\underset{G}{\Longrightarrow}}=\left(\underset{G}{\Rightarrow}\right)^*$ $\overset{+}{\underset{G}{\Longrightarrow}}=\left(\underset{G}{\Rightarrow}\right)^+$

定义 1.4 对于语言 $G=(V,T,P,S)$：
语法范畴 A $L(A)=\left\{w|w\in T^*且A\overset{*}{\Rightarrow}w\right\}$
语言（Language）$L(G)=\left\{w|w\in T^*且S\overset{*}{\Rightarrow}w\right\}$
句子（Sentence）$\forall w\in L(G)$
句型（Sentential Form）$\forall\alpha\in(V\cup T)^*$，如果 $S\overset{*}{\Rightarrow}\alpha$，则称 α 是 G 产生的一个句型。

定义 1.5 对于文法 G_1,G_2，如果 $L(G_1)=L(G_2)$，则称 G_1 与 G_2 等价。

1.1 文法的乔姆斯体系

定义 1.6 对于文法 $G=(V,T,P,S)$：
G 叫做 0 型文法，也叫短语结构文法（PSG，Phrase Structure Grammar）
$L(G)$ 是 0 型语言，也叫短结构语言，可递归枚举集。

定义 1.7 对于 0 型文法文法 $G=(V,T,P,S)$：
$\forall\alpha\to\beta\in P$，$|\beta|\geq|\alpha|$，则 G 是 1 型文法，或上下文有关文法。

定义 1.8 对于 1 型文法文法 $G=(V,T,P,S)$：
$\forall\alpha\to\beta\in P$，$|\beta|\geq|\alpha|$，$\alpha\in V$ 则 G 是 2 型上下文无关文法。

定义 1.9 对于 2 型文法文法 $G=(V,T,P,S)$：$\forall\alpha\to\beta\in P$：$A\to wB$ 和 $A\to w$，其中 $A,B\in V,w\in T^+$：G 是右线性文法。$A\to Bw$ 和 $A\to w$，其中 $A,B\in V,w\in T^+$：G 是左线性文法。则 G 是 3 型文法，或正则文法。

1.2 空产生式

允许在 CSG，CFG，RG 文法中存在空产生式。
允许在 CSL，CFL，RL 语言中存在空语句。
左右线性文法等价，其中左线性的表述好。**左右都有的线性文法不是正则文法。**
$\forall G,\exists G',L(G')=L(G)$，但是 G' 中的开始符号不出现在任何产生式的右部，且在 $\epsilon\in L(G')$ 时，G' 中只有 $S'\to\epsilon$ 这样一个 ϵ 产生式。**可以去 S 产生式和 ϵ 产生式**

1.2.1 语言运算

给定上下文无关文法 G_1,G_2，构造 G 使得：

1. $L(G)=L(G_1)L(G_2)$
   其中 $V_1\cup V_2=\emptyset$，$S\notin V_1\cup V_2$
   $G=(V_1\cup V_2\cup S,T,P_1\cup P_2\cup S\to S_1S_2,S)$
2. $L(G)=L(G_1)\cup L(G_2)$
   $G_\cup=(V_1\cup V_2\cup S,T,P_1\cup P_2\cup P_3,S)$
   $P_3=\{S\to S_1|S_2\}$
3. $L(G)=L(G_1)^*$
   $G_*=(V_1\cup\{S\},T,P_1\cup P_2,S)$
   $P_2=\{S\to\epsilon|SS_1\}$

给定 RG G_1,G_2，构造 RG G 使得：

1. $L(G)=L(G_1)L(G_2)$
   其中 $V_1\cup V_2=\emptyset$，$S\notin V_1\cup V_2$
   $G=(V_1\cup V_2\cup S,T,P_1'\cup P_2\cup P_3,S_1)$
   $P_1'=\{A\to aB|A\to aB\in P_1\}$ $P_3=\{A\to aS_2|A\to a\in P_1\}$
2. $L(G)=L(G_1)\cup L(G_2)$
   $G_\cup=(V_1\cup V_2\cup S,T,P_1\cup P_2\cup P,S)$
   $P=\{S\to\alpha|S_1\to\alpha\in P_1\}\cup\{S\to\alpha|S_2\to\alpha\in P_2\}$
3. $L(G)=L(G_1)^*$
   $G_\cup=(V_1\cup S,T,P_1\cup P,S)$ $P=\{S\to\epsilon|S_1S\}$

定义 1.10 $M=(Q,\Sigma,\delta,q_0,F)$，其中：
Q ：状态的有穷集合
Σ ：输入字母表
δ ：状态转义函数 $Q\times\Sigma\to Q$。$\forall(q,a)\in Q\times\Sigma$，$\delta(q,a)=p$
q_0 ：开始状态
F ：终止状态

定义 1.11 扩展 δ 为 $\hat{\delta}:Q\times\Sigma^*\to Q$

1. $\hat{\delta}(q,\epsilon)=q$
2. $\hat{\delta}(q,wa)=\delta(\hat{\delta}(q,w),a)$

注意到 $Q\subset Q\times\Sigma^*$，且对 $\forall(q,a)\in Q\times\Sigma$，$\hat{\delta}(q,a)=\delta(q,a)$ 所以，不用区分 δ 与 $\hat{\delta}$。
δ 是 $Q\times\Sigma^*\to Q$ 上的映射。

定义 1.12 有穷状态自动机 $M=(Q,\Sigma,\delta,q_0,F)$ 识别的语言：$L(M)=\{x|\delta(q_0,x)\in F\}$

定义 1.13 有穷状态自动机 M_1,M_2：若有 $L(M_1)=L(M_2)$，则称二者等价。

定义 1.14 对于有穷自动状态机 $M=(Q,\Sigma,\delta,q_0,F)$
$set(q)=\{x|\delta(q_0,x)=q\}$
所以有：$\Sigma^*=\bigcup_{q\in Q}set(q)$
$\forall q,p\in Q,q\neq p,set(q)\cap set(p)=\emptyset$
同时，如果 Q 中不存在不可达状态，则 $set(q_0),set(q_1)\dots$ 是 Σ^* 的一个划分。

定义 1.15 $xR_My\iff\delta(q_0,x)=\delta(q_0,y)\iff\exists q\in Q,x,y\in set(q)$

例 1.1 对于 FSA $M=(Q,\Sigma,\delta,q_0,F)$，构造 FA M' 使得 $L(M')=\Sigma^*-L(M)$ $M'=(Q,\Sigma,\delta,q_0,\complement_QF)$

例 1.2 给定 $M_1=(Q_1,\Sigma,\delta_1,q_{01},F_1)$ $M_2=(Q_2,\Sigma,\delta_2,q_{02},F_2)$ 构造：

- M_3 使得 $L(M_3)=L(M_1)\cap L(M_2)$
  $M=(Q_1\times Q_2,\Sigma,\delta,[q_{01},q_{02}],F_1\times F_2)$
- M_3 使得 $L(M_3)=L(M_1)\cup L(M_2)$
  $M=(Q_1\times Q_2,\Sigma,\delta,[q_{01},q_{02}],Q_1\times F_2\cup F_1\times Q_2)$
  $\delta([q,p],a)=[\delta_1(q,a),\delta_2(p,a)]$

1.3 NFA 不确定的有穷状态自动机

定义 1.16 NFA $M=(Q,\Sigma,\delta,q_0,F)$，其中：
Q,Σ,q_0,F 同 FA。
δ: $Q\times\Sigma\to2^Q$（幂集），$\delta(q,a)=\{p_1,p_2,\dots,p_n\}$ 可选一个进入
扩展 δ 为 $\hat{\delta}:Q\times\Sigma^*\to2^Q$ $\hat{\delta}(q,\epsilon)=\{q\}$ $\hat{\delta}(q,wa)=\bigcup_{q_0\in\hat{\delta}(q,w)}\delta(q_0,a)$
进一步拓展：$\hat{\delta}:2^Q\times\Sigma^*\to2^Q$ $\hat{\delta}(P,x)=\bigcup_{p\in P}\hat{\delta}(p,x)$
对 $\forall(q,a)\in Q\times\Sigma^*$，$\hat{\delta}(q,a)=\delta(q,a)$ 所以不用区分 δ 与 $\hat{\delta}$。
δ 是 $Q\times\Sigma^*\to2^Q$ 上的映射。

定理 1.1 NFA 与 DFA 等价。
证：对于 NFA $M=(Q,\Sigma,\delta,q_0,F)$，构造 DFA $M'=(2^Q,\Sigma,\delta',[q_0],F')$
$F'=\{[q_1,q_2,\dots,q_n]|\{q_1,q_2,\dots,q_n\}\cap F\neq\emptyset\}$
$\delta'([q_1,q_2,\dots,q_n],a)=[\beta_1,\beta_2,\dots,\beta_n]\iff\delta(\{q_1,q_2,\dots,q_n\},a)=\{\beta_1,\beta_2,\dots,\beta_n\}$
所以，对于 $\forall M$ 是 NFA，可以构造 DFA M'

1.4 带空移动的有穷状态自动机

定义 1.17 $\epsilon-$NFA $M=(Q,\Sigma,\delta,q_0,F)$：
δ：状态转义函数。$\delta:Q\times(\Sigma\cup\{\epsilon\})\to2^Q$
其中，对于 $\forall q\in Q,\delta(q,\epsilon)=\{p_1,p_2,\dots,p_n\}$：表示在 q 状态不读入任何字符，可以将状态变为 $p_1,p_2,\dots,p_n$，称为 M 在 q 状态做了一次空移动

1. $\epsilon-CLOSURE(q)=\{p|从 q 到 p 有一条标记为\epsilon的路\}$
2. $\epsilon-CLOSURE(P)=\bigcup_{p\in P}\epsilon-CLOSURE(p)$
3. $\hat{\delta}(q,\epsilon)=\epsilon-CLOSURE(q)$
4. $\hat{\delta}(q,wa)=\epsilon-CLOSURE(P)$ $P=\bigcup_{r\in\hat{\delta}(q,w)}\delta(r,a)$
5. 进一步拓展 $\delta:2^Q\times\Sigma\to w^Q$： $\delta(P,a)=\bigcup_{q\in P}\delta(q,a)$
6. 进一步拓展 $\hat{\delta}:2^Q\times\Sigma^*\to2^Q$： $\delta(P,w)=\bigcup_{q\in P}\hat{\delta}(q,w)$
   注意：$\hat{\delta}\neq\delta$，$\delta(q,a)$ 只要 a 不是 ϵ，就不能走空移动。反过来，$\hat{\delta}(q,w)$ 在 w 字符串的中间可以走任意个空移动

定义 1.18 $\epsilon-NFA$ M 识别的语言：$L(M)=\{x|\hat{\delta}(q_0,x)\cap F\neq\emptyset\}$

定理 1.2 $\epsilon-NFA$ 等价与 NFA。
证明：设 $\epsilon-NFA$ $M=(Q,\Sigma,\delta,q_0,F)$：
取 $M'=\left(Q,\Sigma,\hat{\delta},q_0,\begin{cases}F\cup\{q_0\} & F\cap\epsilon-CLOSURE(q_0)\neq\emptyset\\F & F\cap\epsilon-CLOSURE(q_0)=\emptyset\end{cases}\right)$
由此可见，DFA，NFA，$\epsilon-NFA$ 等价。以后统称为 FA。

1.5 FA 与 RG 等价

定理 1.3 对 $\forall$ DFA M，$\exists$ RG G 使得 $L(G)=L(M)$。
构造：对于 DFA $M=(Q,\Sigma,\delta,q_0,F)$，取 RG $G=(Q,\Sigma,P,q_0)$
$P=\{q\to ap|\delta(q,a)=p\}\cup\{q\to a|\forall\delta(q,a)\in F\}\cup\{q_0\to\epsilon|q_0\in F\}$
$=\{q\to ap|\delta(q,a)=p\}\cup\{q\to\epsilon|q\in F\}$

定理 1.4 对 $\forall$ RG G，$\exists$ FA M 使得 $L(G)=L(M)$。
构造：对于 RG $G=(V,T,P,S)$，RG 为右线性文法。
取 FA $M=(V\cup\{f\},T,\delta,S,\{f\})$，其中
$\delta(A,a)=\{B|\forall A\to aB\in P\}\cup\{f|\forall A\to a\in P\}$

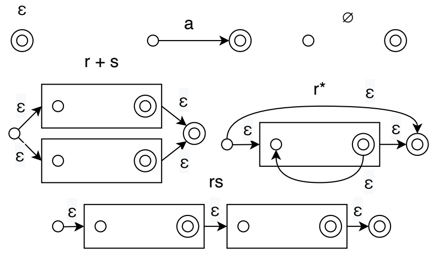
2 正则表达式

定义 2.1 σ 上的 RE 是满足如下条件的式子：

- ϵ 是 RE，表示的语言 $L(\epsilon)=\{\epsilon\}$
- $\emptyset$ 是 RE，表示的语言 $L(\emptyset)=\emptyset$
- $\forall a\in\sigma$，a 是 RE，表示的语言 $L(a)=\{a\}$
- $\forall r,s$ 是 RE 且分别表示 R,S，则有：
  - $(r+s)$ 是 RE，表达的语言 $L(r+s)=R\cup S$
  - (rs) 是 RE，表达的语言 $L(r+s)=RS$
  - (r^*) 是 RE，表达的语言 $L(r+s)=R^*$

约定 1：运算符优先级 $*>\times>+$
约定 2：引入正闭包 $r^+=r^*r$

2.1 RE →FA



定理 2.1 对于 $\forall$ RE r，$\exists$ FA M，使得 $L(M)=L(r)$

3 正则语言的性质

3.1 正则语言的泵引理

引理 3.1 正则语言的的泵引理。对于 $\forall$ RL L，存在一个仅依赖与 L 的正整数 N，对于 $\forall z \in L, |z| \geq N$，则存在 u, v, w 满足以下条件：

$$uvw = z \quad |uv| \leq N \quad |v| \geq 1 \quad \text{对于} \ \forall k \geq 0,\ uv^k w \in L$$

引理 3.2 拓展的泵引理。对于 $\forall$ RL L，存在一个仅依赖与 L 的正整数 N，对于 $\forall z = z_1 z_2 z_3 \in L, |z_2| \geq N$，则存在 u, v, w 满足以下条件：

$$uvw = z_2 \quad |uv| \leq N \quad |v| \geq 1 \quad \text{对于} \forall k \geq 0, uv^k w \in L$$

例 3.1 证明 $L = \{0^p | p\text{是质数}\}$ 不是 RL。
设 L 是 RL。N 是泵引理所说正整数。设 x 为大于 N 的最小质数。
取 $z = 0^x$。取 $v = 0^l$，$l \geq 1$。$uv^k w = 0^{x+(k-1)l}$。
当 $k = x + 1$ 时，有 $x + (x + 1 - 1)l = (l + 1)x$ 是合数，所以 $uv^k w \notin L$，与泵引理矛盾。所以 L 不是 RL。

定理 3.1 RL 对于并、乘、闭包、交、补封闭。

3.2 Myhill 定理

定义 3.1 $xR_M y \iff \delta(q_0, x) = \delta(q_0, y)$

$$\iff \exists q \in Q, x, y \in set(q)$$

定义 3.2 $xR_L y \iff \forall z \in \Sigma^*, xz \in L \iff yz \in L$

定义 3.3 R 是右不变的是指如果 xRy 则有 $\forall z \in \Sigma^*, xzRyz$

定理 3.2 对于 DFA M，如果 $xR_M y$，则 $xR_{L(M)} y$。
证明：由 $xR_M y$ 知 $\delta(q_0, x) = \delta(q_0, y)$。不妨设其为 q_1。
对于 $\forall z \in \Sigma^*$，$\delta(q_0, xz) = \delta(\delta(q_0, x), z) = \delta(q_1, z) = \delta(q_0, yz)$。不妨设其为 q_2。
如果 $xz \in L$，则有 $q_2 \in F$，所以 $yz \in L$。反之亦然。
所以 $xz \in L \iff yz \in L$。所以 $xR_M y \iff xR_{L(M)} y$ □

定理 3.3 Myhill 定理。以下 3 个命题等价：

1. L 是 RL
2. L 是 Σ^* 上的某个具有有穷指数的右不变等价关系的某些等价类的并。
3. R_L 具有有穷指数。

证明 3.1 $1 \Rightarrow 2$。设 L 是 RL，DFA M 使得 $L(M) = L$。
R_M 是 Σ^* 上的右不变等价关系，且 $|\Sigma^*/R_M| \leq |Q|$。
$L = \bigcup_{q \in F} set(q)$

证明 3.2 $2 \Rightarrow 3$。设 L 是 Σ^* 上的具有有穷指数的右不变等价关系 R 的某些等价类的并。
下面证明如果 xRy，那么 $xR_L y$。
设 $x, y \in \Sigma^*, xRy$。由 R 的右不变性，对于 $\forall z \in \Sigma^*$，$xzRyz$。
再注意到 L 是 R 的某些等价类的并，且 xz, yz 在同一个等价类中，所以 $xz \in L \iff yz \in L$。
所以 $xR_L y$。

证明 3.3 $3 \Rightarrow 1$。设 R_L 具有有穷指数。取 $M' = (\Sigma^*/R_L, \Sigma, \delta', [\epsilon], \{[x] | x \in L\})$
$\delta'([a], x) = [ax]$。显然，$L(M') = L$。

意义：L 是 RL 的充要条件

定理 3.4 在同构意义下，M' 是状态最少的唯一识别 L 的 DFA。

3.2.1 证明 L 是/不是 RL

1. 证明 L 是 RL。证明 R_L 的指数有穷。**列举所有等价类即可**
2. 证明 L 不是 RL。证明 R_L 的指数无穷。

例 3.2 证明 $\{0^n 1^n | n \geq 0\}$ 不是 RL。
易得，对于 $\forall i \neq j \in N$，$0^i \not{R_L} 0^j$ 因此 R_L 的指数无穷。

3.3 DFA 的最小化

标记终止状态和非终止状态不能合并
逐个判断其他的能否合并

4 CFL 上下文无关语言

4.1 CFG

定义 4.1 CFG $G = (V, T, P, S)$ 的语法树为满足如下条件的树：

1. 每个节点的标记 $x \in V \cup T \cup \{\epsilon\}$
2. 如果节点 V 的标记为 A，V 从左到右的子节点 $V_1, V_2, \ldots, V_k$ 的标记依次为 $y_1, y_2, \ldots, y_k$，则 $A \to y_1 y_2 \ldots y_k \in P$
3. 根节点标记为 S
4. 中间节点的标记为变量 $x \in V$
5. 从左到右的叶子节点 $v_1, \ldots, v_n$ 的标记 $x_1, \ldots, x_n$ 组成的串 $x_1 x_2 \ldots x_n$ 为该树的结果。
6. 如果 v 的标记为 ϵ，则它没有兄弟。

定义 4.2 满足语法树定义中除第三条外条件的树，称作 A-子树。

定理 4.1 有一颗结果为 α 的语法树 $\iff S \overset{*}{\Rightarrow} \alpha$

定义 4.3 每一步派生均实施在当前句型最右变量上的派生叫最右派生。
每一步派生均实施在当前句型最左变量上的派生叫最左派生。

定理 4.2 最左派生与最右派生的语法树是一一对应的。

定义 4.4 如果 CFG G 有句子有棵颗不同的语法书，则 G 是二义性的。

定义 4.5 如果 CFL L 没有非二义性文法，则称之为固有二义性的。

4.2 去无用符号

4.2.1 去除无用符号

定义 4.6 X 是有用符号，即 $\exists X \in L(G)$，$S \overset{*}{\Rightarrow} \alpha X \beta \overset{*}{\Rightarrow} x$。
X 是有用的，必须同时满足如下两条：

1. $S \overset{*}{\Rightarrow} \alpha X \beta$
   如何判断？
   (a) $V' = \{S\} \cup \{A | S \to \alpha A \beta \in P\}$
       $T' = \{a | S \to \alpha a \beta \in P\}$
   (b) 重复
       $V' = V' \cup \{B | A \to \alpha B \beta \in P \& A in V'\}$
       $T' = T' \cup \{a | A \to \alpha a \beta \in P \& A in V'\}$
2. $X \overset{*}{\Rightarrow} w, w \in T^*$
   $G = (V, T, P, S)$，对于：
   $\forall a \in T$，a 满足 2。
   $\forall A \in V$，如何判断 A 是否满足 2?
   (a) $V' = \{A | A \to w \in P\}$
   (b) 重复 $V' = \{A | A \to \alpha, \alpha \in (V' \cup T)^*\} \cup V'$

定理 4.3 对于 $\forall$ CFG G，$\exists$ CFG G'，

1. $L(G') = L(G)$
2. G' 中无无用符号。

4.2.2 去除 $\epsilon-$ 产生式

定义 4.7 如果 $A \overset{*}{\Rightarrow} \epsilon$，则称 A 为可空变量。

定理 4.4 对于 $\forall$ CFG G, G'：

1. G' 中无空产生式
2. $L(G') = L(G) - \{\epsilon\}$

对于 $A \to x_1 x_2 \ldots x_n$，替换为 $A \to y_1 y_2 \ldots y_n$，其中当 x_i 不是可空变量时，$y_i = x_i$，否则 $y_i = x_i$ 或 ϵ。
注意 $y_1 y_2 \ldots y_n$ 不能都为 ϵ。

4.2.3 去单一产生式

定义 4.8 形如 $A \to B$ 的产生式是单一产生式。

定义 4.9 对 $\forall$ CFG G，$\exists$ CFG $G', L(G') = L(G)$，G' 中无单一产生式。
推论：对于 $\forall$ CFG G，存在 CFG G' 使得 $L(G') = L(G)$，且 G' 没有无用符，$\epsilon-$ 产生式和单一产生式。

4.3 CNF 乔姆斯基范式

定义 4.10 如果 G 的产生式均具有以下形式：

$$\begin{cases} A \to BC \\ A \to a \end{cases}$$

则称之为 CNF。

4.4 GNF 格雷巴赫范式

定义 4.11 如果 G 的产生式均具有以下形式：
$A \to a\alpha, \alpha \in V^*$
则称之为 GNF。

例 4.1 $A \to Aa|Ab|c|d = \begin{cases} A \to c|d|cB|dB \\ B \to a|b|aB|bB \end{cases}$

例 4.2 $A \to A\alpha_1|A\alpha_2|A\alpha_3|\beta_1|\beta_2 = \begin{cases} A \to \beta_1|\beta_2|\beta_1 B|\beta_2 B \\ B \to \alpha_1|\alpha_2|\alpha_3|\alpha_1 B|\alpha_2 B|\alpha_3 B \end{cases}$

例 4.3 $\begin{cases} S \to ABS|BAA \\ A \to BB|BAA \\ B \to aB|b \end{cases} = \begin{cases} S \to aBBBS|bBBS|aBAABS|bAABSaBAA \\ A \to aBB|bB|aBAA|bAA \\ B \to aB|b \end{cases}$

步骤：

1. 给变量排序
2. 从 A_1 到 A_n 逐一使产生式满足如下要求：
   $A_i \to A_h \alpha, j \geq i$
3. 从 A_{n-1} 开始通过回代，逐一使 $A_{n-1}, A_{n-2} \ldots$ 的产生式满足要求
4. 通过代入，使第二步中引入的新变量的产生式满足要求。
   关键：**去左递归**
   如：
   $\begin{cases} A \to A\alpha_1|A\alpha_2|\ldots|A\alpha_n \\ A \to \beta_1|\beta_2|\ldots|\beta_m \end{cases}$ 为所有 A 产生式，且 $\beta_1, \beta_2, \ldots, \beta_m$ 的首字母不是 A，可以用如下的产生式组替代：
   $\begin{cases} A \to \beta_1|\beta_2|\ldots|\beta_m|\beta_1 A'|\beta_2 A'|\ldots|\beta_i A' \\ A' \to \alpha_1 A'|\alpha_2 A'|\ldots|\alpha_n A'|\alpha_1|\alpha_2|\ldots|\alpha_n \end{cases}$

定理 4.5 对于 $\forall$ 化简了的 CFG，$\exists$ GNF 与之等价。

5 PDA 下推自动机

5.1 PDA 的定义

定义 5.1 PDA $M = (Q, \Sigma, \Gamma, \delta, q_0, z_0, F)$ 其中：

Q ： 状态的有穷集合

Σ ： 输入字母表

Γ ： 栈符号的非空有穷集

δ ： 状态转义函数
$Q \times (\Sigma \cup \{\epsilon\}) \times \Gamma \to 2^{Q \times \Gamma^*}$。
$\forall (q, a, A) \in Q \times (\Sigma \cup \{\epsilon\}) \times \Gamma$，$\delta(q, a, A) = \{(p_1, \gamma_1), \ldots, (p_k, r_k)\}$ 表示 M 在状态 q，栈顶为 A 时读到 a，将栈顶符号 A 弹出，将 γ_i 依次压入栈，并将状态改为 p_i。
$\forall (q, A) \in Q \times \Gamma$，$\delta(q, \epsilon, A) = \{(p_1, \gamma_1), \ldots, (p_k, r_k)\}$ 表示 M 在状态 q，栈顶为 A 时做空移动，将栈顶符号 A 弹出，将 γ_i 依次压入栈，并将状态改为 p_i。

q_0 ： 开始状态

$z_0 \in \Gamma$ ： 栈底符号

F ： 终止状态

定义 5.2 PDA M 的 ID $(q, x, \alpha) = Q \times \Sigma^* \times \Gamma^*$
其中 q 是 M 的当前状态，x 是 M 的输入带上剩余串，α 是 M 的栈中当前的内容。
设 M 当前的 ID 是 $(q, ax, A\alpha)$，如果 $(p, \gamma) \in \delta(q, a, A)$，则 M 的 ID 变为 $(p, x, \gamma\alpha)$，记作 $(q, ax, A\alpha) \vdash_M (p, x, \gamma\alpha)$。
如果 $(p, \gamma) \in \delta(q, \epsilon, A)$，则 M 的 ID 变为 $(p, ax, \gamma\alpha)$，记作 $(q, ax, A\alpha) \vdash_M (p, ax, \gamma\alpha)$。
$\vdash_M$ 是 $Q \times \Sigma^* \times \Gamma^*$ 上的二元关系。

定义 5.3 $M = (Q, \Sigma, \Gamma, \delta, q_0, z_0, F)$
用终态识别的语言 $L(M) = x | (q_0, x, z_0) \vdash^* (q, \epsilon, \alpha)$ and $q \in F$。
用空栈识别的语言 $N(M) = x | (q_0, x, z_0) \vdash^* (q, \epsilon, \epsilon)$ and $q \in F$。