编译作业参考答案 第三/四次作业

P56 1

画出下列文法的状态图

```
<Z> ::= <B>e
<B> ::= <A>f
<A> ::= e | <A>e
```

并使用该状态图检查下列句子，是否是该文法的合法句子：`f`，`eeff`，`eefe`

答案解析

本题的状态图是唯一的：

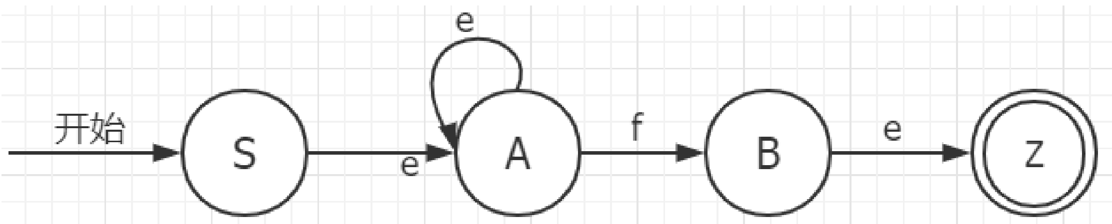

Figure 1: P56.1 状态图

`f` 和 `eeff` 不是，`eefe` 是

批改标准

- 状态图画对
- 三个串判断正确

P56 2

题目给出状态图，其中 S 为始态，Z 为终态：

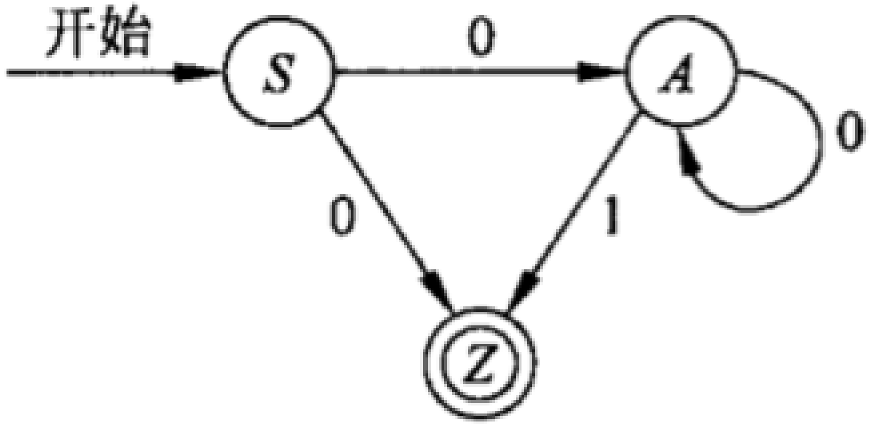

Figure 2: P56.2 图

（1）写出相应的正则文法

（2）写出该文法的 V, V_n, V_t

（3）该文法确定的语言是什么？

答案解析

文法：

```
<Z> ::= 0 | <A>1
<A> ::= 0 | <A>0
```

$V_n = \{< Z >, < A >\}, V_t = \{0, 1\}, V = V_n \cup V_t$

由于`<A> ::= 0|<A>0`，改写为 EBNF 可以得到`<A> ::= 0{0}`;

所以`<Z> ::= 0 | 0{0}1`，因此语言是：$\{0\} \cup \{0^n 1 \mid n \geq 1\}$（使用文字描述也可以）

批改标准

- 文法和参考答案等价
- 三个集合和自己的文法对应且正确
- 语言和参考答案等价

P76 1

令 A、B 和 C 是任意正则表达式，证明以下关系成立：

- $A|A = A$
- $(A^*)^* = A^*$
- $A^* = \epsilon|AA^*$
- $(AB)^*A = A(BA)^*$
- $(A|B)^* = (A^*B^*)^* = (A^*|B^*)^*$

答案解析

很多东西一看看上去很显然，但是为了严谨性，以下全都通过证明等号两边的正则式定义的语言相同来证明正则式等价。

证明的基础是集合运算，请注意集合的笛卡尔乘积**不具有交换性**，即对于任意集合 A, B，不一定有 $AB = BA$ 成立。但是可以找到一些特例，比如对于串的运算，如果 $A = \{\epsilon\}$，那么对于任意 B，都有 $AB = BA$。

使用 r 代表正则式，$L(r)$ 代表其定义的语言（语言是一个集合）：

- 连接，$L(r_1r_2) = \{xy \mid x \in L(r_1), y \in L(r_2)\} = L(r_1)L(r_2)$

- 选择，$L(r_1 \mid r_2) = \{x \mid x \in L(r_1) \vee x \in L(r_2)\} = L(r_1) \cup L(r_2)$
- 重复，$L(r^*) = \bigcup\limits_{i=0}^{\infty} L^i(r)$

为了不那么辣眼睛，**以下均简写** $L_A = L(A)$

(1) $A|A = A$

$$
\begin{aligned}
&\because L_{A|A} = L_A \cup L_A = L_A \\
&\therefore A|A = A
\end{aligned}
$$

(2) $(A^*)^* = A^*$

使用数学归纳法易证得：对于集合 S，令 $U = \bigcup\limits_{i=0}^{\infty} S^i$，有 $U^k = U, k \in^+ :$

$k = 1$ 时，$U^k = U$ 显然成立。

设 $k = n$ 时，$U^k = U^n = U$ 成立，则：

$$
\begin{aligned}
U^{n+1} &= U^n U = UU = U^2 \\
&= (\bigcup_{i=0}^{\infty} S^i)(\bigcup_{j=0}^{\infty} S^j) \\
&= \bigcup_{i=0}^{\infty} (S^i \bigcup_{j=0}^{\infty} S^j) \\
&= \bigcup_{i=0}^{\infty} (\bigcup_{j=0}^{\infty} S^{i+j}) \\
&= \bigcup_{i=0}^{\infty} (\bigcup_{j=i}^{\infty} S^j) \\
&= \bigcup_{j=0}^{\infty} S^j \\
&= U
\end{aligned}
$$

证毕。因此令 $S = L_A, U = \bigcup\limits_{i=0}^{\infty} L_A^i = L_{A^*}$，则有:

$$L_{(A^*)^*} = \bigcup_{i=0}^{\infty} L_{A^*}^i = \bigcup_{i=0}^{\infty} U^i = \bigcup_{i=0}^{\infty} U = U = L_{A^*} \therefore (A^*)^* = A^*$$

(3) $A^* = \epsilon | AA^*$

$$\begin{aligned}
\because L_{\epsilon|AA^*} &= L_\epsilon \cup L_{AA^*} \\
&= L_\epsilon \cup L_A L_{A^*} \\
&= L_A^0 \cup (L_A \bigcup_{i=0}^{\infty} L_A^i) \\
&= L_A^0 \cup \bigcup_{i=1}^{\infty} L_A^i \\
&= \bigcup_{i=0}^{\infty} L_A^i \\
&= L_{A^*} \\
\therefore A^* &= \epsilon | AA^*
\end{aligned}$$

(4) $(AB)^* A = A(BA)^*$

使用数学归纳法证明，对于集合 S_1, S_2，有 $(S_1 S_2)^k S_1 = S_1 (S_2 S_1)^k, k \in$:

- $k = 0$ 时，显然成立
- 设 $k = n$ 时成立，即 $(S_1 S_2)^n S_1 = S_1 (S_2 S_1)^n$，则 $(S_1 S_2)^{n+1} S_1 = (S_1 S_2)^n S_1 S_2 S_1 = S_1 (S_2 S_1)^n S_2 S_1 = S_1 (S_2 S_1)^{n+1}$

证毕。则：

$$
\begin{aligned}
L_{(AB)^*A} &= L_{(AB)^*}L_A \\
&= (\bigcup_{i=0}^{\infty} L_{AB}^i)(L_A) \\
&= (\bigcup_{i=0}^{\infty} (L_A L_B)^i)(L_A) \\
&= \bigcup_{i=0}^{\infty} (L_A L_B)^i L_A \\
&= \bigcup_{i=0}^{\infty} L_A (L_B L_A)^i \\
&= L_A \bigcup_{i=0}^{\infty} (L_B L_A)^i \\
&= L_A \bigcup_{i=0}^{\infty} L_{BA}^i \\
&= L_A L_{(BA)^*} \\
&= L_{A(BA)^*} \\
\therefore (AB)^*A &= A(BA)^*
\end{aligned}
$$

(5) $(A|B)^* = (A^*B^*)^* = (A^*|B^*)^*$

思路 1（工作量较大）

使用的证明工具是定理：若集合 $S_1 \subset S_2, S_2 \subset S_1$，则 $S_1 = S_2$。

证明本题需要一些前置结论：

- 对于集合 S_1, S_2，若 $S_1 \subset S_2$，那么 $\bigcup_{i=0}^{\infty} S_1^i \subset \bigcup_{i=0}^{\infty} S_2^i$
  - 即对于语言 $L_A \subset L_B$，有 $L_{A^*} \subset L_{B^*}$
- 对于集合 S_1, S_2, S_3, S_4，若 $S_1 \subset S_3, S_2 \subset S_4$，那么 $S_1S_2 \subset S_3S_4$
  - 即对于语言 $L_A \subset L_C, L_B \subset L_D$，有 $L_{AB} \subset L_{CD}$

证明下面的 6 个子命题（只要证明出任意 4 个，就可以通过 $\subset$ 的传递性得到剩下的 2 个；或者证明 3 个构成环形 $\subset$ 关系的子命题，也能通过传递性得到剩下的 3 个）：

1. 证明 $L_{(A|B)^*} \subset L_{(A^*|B^*)^*}$：

$$
\begin{aligned}
\because L_A &\subset \bigcup_{i=0}^{\infty} L_A^i = L_{A^*} \\
L_B &\subset \bigcup_{j=0}^{\infty} L_B^j = L_{B^*} \\
\therefore L_{A|B} &= L_A \cup L_B \subset L_{A^*} \cup L_{B^*} = L_{A^*|B^*} \\
\therefore L_{(A|B)^*} &\subset L_{(A^*|B^*)^*}
\end{aligned}
$$

2. 证明 $L_{(A|B)^*} \subset L_{(A^*B^*)^*}$：

$$
\begin{aligned}
\because L_A, L_A^0 &\subset \bigcup_{i=0}^{\infty} L_A^i = L_{A^*} \\
L_B, L_B^0 &\subset \bigcup_{j=0}^{\infty} L_B^j = L_{B^*} \\
\therefore L_A L_B^0, L_A^0 L_B &\subset L_{A^*} L_{B^*} = L_{A^*B^*} \\
\therefore L_{A|B} &= L_A \cup L_B = L_A L_B^0 \cup L_A^0 L_B \subset L_{A^*B^*} \\
\therefore L_{(A|B)^*} &\subset L_{(A^*B^*)^*}
\end{aligned}
$$

3. 证明 $L_{(A^*|B^*)^*} \subset L_{(A|B)^*}$：

   由之前的证明题 (2)，可知 $(A|B)^* = ((A|B)^*)^*$

$$
\begin{aligned}
&\because L_A, L_B \subset L_A \cup L_B = L_{A|B} \\
&\therefore L_{A^*}, L_{B^*} \subset L_{(A|B)^*} \\
&\therefore L_{A^*|B^*} = L_{A^*} \cup L_{B^*} \subset L_{(A|B)^*} \\
&\therefore L_{(A^*|B^*)^*} \subset L_{((A|B)^*)^*} = L_{(A|B)^*}
\end{aligned}
$$

4. 证明 $L_{(A^*|B^*)^*} \subset L_{(A^*B^*)^*}$：

$$\because L_A^0 \subset L_{A^*}$$
$$L_B^0 \subset L_{B^*}$$
$$\therefore L_{A^*} = L_{A^*}L_B^0 \subset L_{A^*B^*}$$
$$L_{B^*} = L_A^0 L_{B^*} \subset L_{A^*B^*}$$
$$\therefore L_{A^*|B^*} = L_{A^*} \cup L_{B^*} \subset L_{A^*B^*}$$
$$\therefore L_{(A^*|B^*)^*} \subset L_{(A^*B^*)^*}$$

5. 证明 $L_{(A^*B^*)^*} \subset L_{(A^*|B^*)^*}$：

$$\because L_{A^*}, L_{B^*} \subset L_{A^*} \cup L_{B^*} \subset L_{A^*|B^*}$$
$$\therefore L_{A^*B^*} = L_{A^*}L_{B^*} \subset L_{A^*|B^*}L_{A^*|B^*} = L_{(A^*|B^*)(A^*|B^*)} \subset L_{(A^*|B^*)^*}$$
$$\therefore L_{(A^*B^*)^*} \subset L_{((A^*|B^*)^*)^*} = L_{(A^*|B^*)^*}$$

6. 证明 $L_{(A^*B^*)^*} \subset L_{(A|B)^*}$：

$$\because L_A, L_B \subset L_A \cup L_B \subset L_{A|B}$$
$$\therefore L_{A^*}, L_{B^*} \subset L_{(A|B)^*}$$
$$\therefore L_{A^*B^*} = L_{A^*}L_{B^*} \subset L_{(A|B)^*}L_{(A|B)^*} = L_{(A|B)^*(A|B)^*} \subset L_{((A|B)^*)^*} = L_{(A|B)^*}$$
$$\therefore L_{(A^*B^*)^*} \subset L_{((A|B)^*)^*} = L_{(A|B)^*}$$

由上述 6 个子命题，可以得到 $L_{(A|B)^*} = L_{(A^*B^*)^*} = L_{(A^*|B^*)^*}$；

所以 $(A|B)^* = (A^*B^*)^* = (A^*|B^*)^*$。

思路 2（工作量较大）

在以下 3 组中证明出来任意 2 组：

- 使用思路 1 中的 1, 3 证明 $(A|B)^* = (A^*|B^*)^*$
- 使用思路 1 中的 2, 6 证明 $(A|B)^* = (A^*B^*)^*$
- 使用思路 1 中的 4, 5 证明 $(A^*|B^*)^* = (A^*B^*)^*$

由 $=$ 的传递性，知 $(A|B)^* = (A^*B^*)^* = (A^*|B^*)^*$。

思路 3（工作量较小）

使用思路 1 中的 2, 6 证明出来 $(A|B)^* = (A^*B^*)^*$;

将 A^*, B^* 代入等式，有 $(A^*|B^*)^* = ((A^*)^*(B^*)^*)^* = (A^*B^*)^*$;

由 $=$ 的传递性，知 $(A|B)^* = (A^*B^*)^* = (A^*|B^*)^*$。

批改标准

证明过程正确就行

学过自动机之后 (5) 的证明方法

助教的碎碎念：为什么上届助教写标答的时候没用自动机呢？

$(A|B)^*$ 、$(A^*B^*)^*$、$(A^*|B^*)^*$ 对应的 NFA 分别如下：

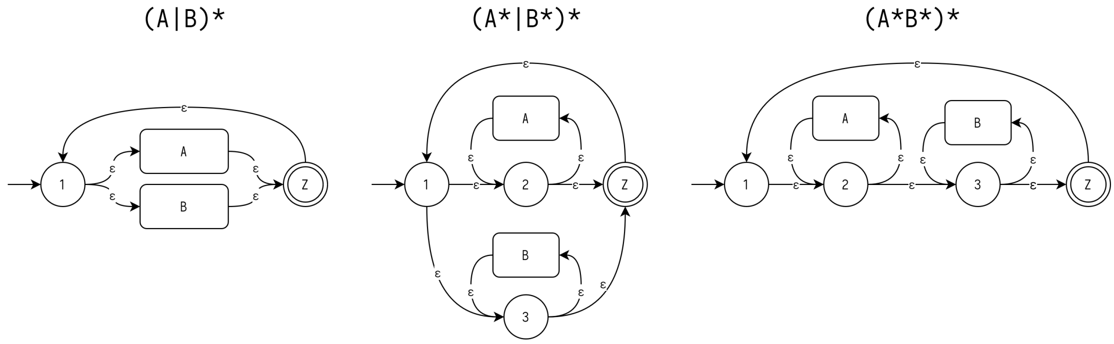

Figure 3: 正则对应的自动机

显然所有露在外面的状态在同一个 ϵ 闭包内，所以最终都可以部分确定化成同样的自动机，如下：

补充习题

证明正则表达式 `a(a|b)*` 和正则文法 $G[Z]$: `Z ::= Za|Zb|a` 等价

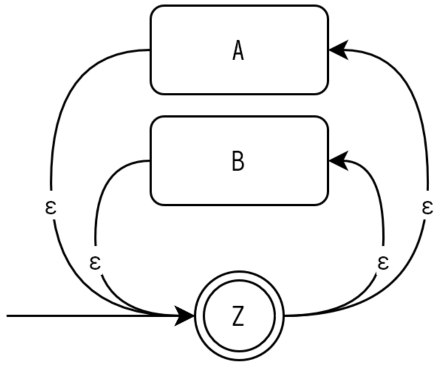

Figure 4: 化简后的自动机

答案解析

自然语言证明

`a(a|b)*` 表示的是以 `a` 打头的串，`a` 的后面可以随意跟 `a` 和 `b`。`Z ::= Za|Zb|a` 定义的句子集合，最短的是 `a`，然后是 `aa`，`ab`，长度为 3 的是 `aaa`、`aab`、`aba`、`abb`，长度为 4 的是……。可以看出也是都以 `a` 打头，后面跟任意个 `a` 或 `b`。两种表示方法所定义的句子集合是一样的，从而证明正则表达式 `a(a|b)*` 和正则文法 `Z ::= Za|Zb|a` 等价。证毕。

通过转换文法为正则表达式来证明

注意书上给出的正则文法转换正则表达式规则中，使用的是右线性的文法，而本题是左线性的，转换时注意不能直接套用。

思路 1:

对 `Z -> Za|Zb|a` 消除 `Za` 带来的递归，得：`Z -> (Zb|a)(a|b)*`，即 `Z -> Zb(a|b)*|a(a|b)*`。

继续消除递归得：`Z -> a(a|b)*(b(a|b)*)*`，即 `Z -> a(a|b)*(a|b)*`，即 `Z -> a(a|b)*` 。

得到了完全一致的正则表达式，无需证明语言等价，证毕。

思路 2:

提取左公因子 `Z`，得 `Z -> Z(a|b)|a`；

消除递归，得：`Z -> a(a|b)*`

得到了完全一致的正则表达式，无需证明语言等价，证毕。