1.画出正规表达式 (a|ba)*相应的确定有限自动机 DFA。（不要求写出构造过程）

（国防科技大学 1996 年硕士生入学考试试题）

分析

这是一类题型，即根据正规表达式来求相应的有限自动机。对于这种类型的题目的固定解法就是：第一步：根据正规表达式画出对应的非确定有限自动机 NFA，即首先对于正规表达式 V，我们将它表示成图 2.1 的拓广转换图，其中结点 X 是初始结点，Y 是唯一的终止结点，正规表达式 V 就是 X、Y 之间的弧的标记。

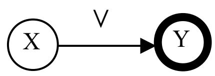

图 2.1 正规表达式 V 的拓广转换图

然后按图 2.2 的转换规则对拓广转换图进行分裂，逐步把这个图转换成每条弧的标记为Σ上的一个字符或 ε，分裂过程中所引进的新结点都是中间结点。显然完全分裂后我们所得到的就是一个非确定有限自动机 M，并且有 L（M）=L（V）。

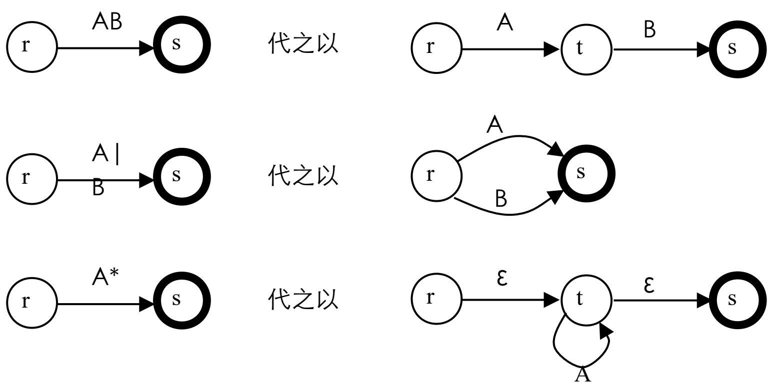

图 2.2 转换规则

第二步：将非确定有限自动机 M 化为题目所要求的形式，一般包括确定化和最小化两个步骤。

确定化的方法如下：

1．定义ε闭包，ε_CLOSURE(I)为

- 若q∈I，则q∈ε_CLOSURE(I)；
- 若q∈I，那么从q出发经任意条ε弧而能到达的任何状态q′都属于ε_CLOSURE(I)；

2．假定I是M的状态集的子集，a∈Σ，定义

$$I_a=\varepsilon\_CLOSURE(J)$$

其中，J是那些可从I中的某一状态结点出发经过一条a弧而到达的状态结点的全体。

3．假定π={a_1,…, a_k}。我们构造一张表，此表的每一行含有k+1列。置该表的首行首列为ε_CLOSURE(X)。一般而言，如果某一行的第一列的状态子集已经确定，例如记为I，那么，置该行的i+1列为I_{ai}（i=1,…, k）。然后，检查该行上的所有状态子集，看它们是否已在表的第一列中出现，将未曾出现者填入到后面空行的第一列。重复上述过程，直至出现在第i+1列（i=1,…, k）上的所有状态子集均已在第一列上出现。因为M的状态子集的个数是有限的，所以上述过程必定在有限步内终止。

4．将构造出来的表视为状态转换表，将其中的每个状态子集视为新的状态。显然，该表唯一地刻划了一个确定有限自动机M′。它的初态是该表首行首列的那个状态，终态是那些含有原终态的状态子集。根据上述构造方法，不难得出：L(M′) = L(M) = L(V)。

DFA的最小化的目的是：为DFA M′寻找一个状态数比它少的DFA M″，并使得L(M″) = L(M′)。

DFA 的最小化的思想是：将 DFA 的状态集分割成一些不相交的子集，使得任何不同的两子集中的状态都是可区别的，而同一子集中的任何两个状态都是等价的。最后，在每个子集中选出一个代表，同时消去其它等价状态。

DFA 的最小化的过程见 2.1 复习提要。

本题的解法为：

第一步：根据正规表达式构造相应的非确定有限自动机 NFA

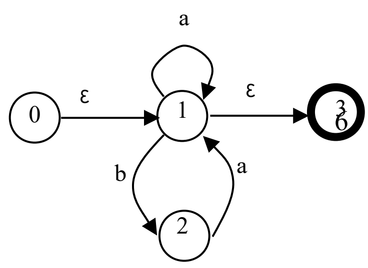

图 2.3 NFA 图

检查所得到的 NFA，看是否是题目所需要的答案，是则结束，不是则进行第二步。

第二步：对图 2.3 中的 NFA 构造相应的状态转换表（NFA 的确定化过程），得到表 2.1：

表 2.1

I	I_a	I_b
{0,1,3}	{1,3}	{2}
{1,3}	{1,3}	{2}
{2}	{1,3}	

从表中可以看出，状态转换图包含 3 个状态子集，{0，1，3}、{1，3}和{2}，我们分别用 3 个不同的状态 0、1、2 来表示它们，这里的状态 0、1、2 和图 2.1 中的状态 0、1、2 不同，它们只是以来表示子集{0，1，3}、{1，3}、{2}的状态名，也可以用其它名字如 i、j、k 来代替。则得到的新的状态转换表为：

表 2.2

I	I_a	I_b

0	1	2
1	1	2
2	1	

在原 NFA 图中状态 0 为初始状态，状态 3 为终止状态，则在新的状态转换表（表 2.2）中结点 0（代表子集{0、1、3}）既是初始状态（包含原初始状态 0），又是终止状态（包含原终止状态 3），新结点 1（代表子集{1、3}）是终止状态（包含原终止状态 3）。根据表 2.2，我们可以得到确定化后的自动机如图 2.4 所示。

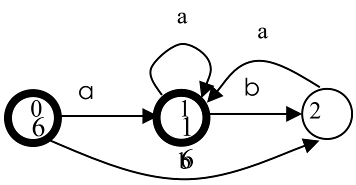

图 2.4 确定化后的 DFA

由于本题所求的就是确定的有限自动机，所以图 2.4 就是本题的正确答案。

如果题目要求最小化的 DFA，则还需要对该 DFA 进行进一步的最小化工作。

最小化的步骤为:

1．将结点 0、1、2 按终止结点和非终止结点分为两个子集{0、1}和{2}，检查子集{0、1}，发现 $I_a^{\{0、1\}}=\{1\}\subseteq\{0，1\}$，$I_b^{\{0、1\}}=\{1\}\subseteq\{2\}$，不需要进行进一步的划分，而子集{2}只包含一个结点，不能进行进一步的划分，所以状态集可以划分为两个子集，{0、1}和{2}，用结点 0 代替子集{0、1}，用结点 1 代替子集{2}，可以得到最小化后的 DFA，如图 2.5。

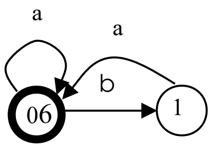

图 2.5 最小化后的 DFA

解答

正规表达式 (a|ba)*相应的确定有限自动机 DFA 如图 2.5 所示。