

2021-2022赛季软191级队学习资料 (2)

编译原理答题攻略

前言

编译原理的历年试卷其实题型十分固定。在做了三年的考题后，我决定为大家总结一下有关编译原理我的一些做题心得。固然，编译原理如果想要得到高分，需要课堂中的不懈努力，但是我认为更加重要的是对于应试技巧的摸索。本资料聚焦编译原理的往年题型，老师划的考试重点以及对于往年真题按照题型的解析，每种题型可能还会给出一些课上习题进行讲解。希望能为级队同学们的复习起到一个抛砖引玉的作用。

题型

编译原理这三年来的题型是这样的：

1. 填空题（每题2分，共10道，总计20分）
2. 选择题（每题2分，共10道，总计20分）
3. 第三题：NFA转化DFA / DFA化简 / 根据语言叙述画出DFA/NFA / 正规式
4. 第四题：给出语法分析推导及语法分析树
5. 第五题：求First和Follow集，画出LL（1）分析表并判断是否为LL（1）文法
6. 第六题：构造识别活前缀DFA, 构造LR或SLR分析表，判断文法是否为LR 或 SLR文法
7. 第七题：根据表达式写出三地址代码或语法树 / 根据程序段写出流图
8. 第八题：根据程序段绘制活动树 / 活动记录 / 控制链 / 访问链
9. 第九题：根据产生式写出语法制导定义和翻译方案

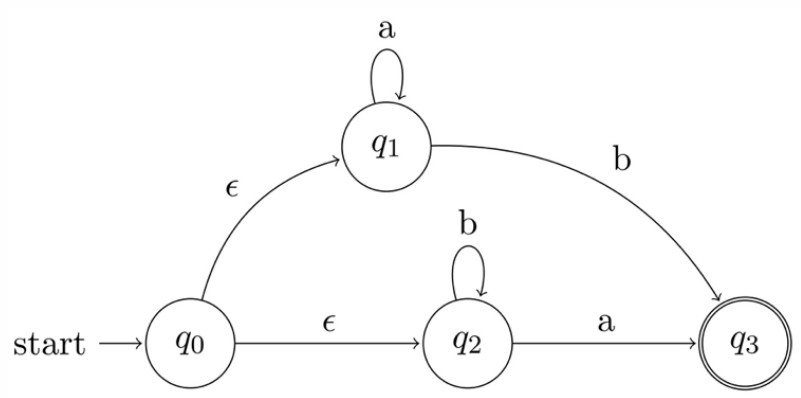
以上题型可能一个题型会出不止一道大题，具体考出哪些大题和这些大题的顺序，我们也只能考场上见分晓了。

题型解析

第三大题

NFA转化DFA

例题1：请将下图的NFA转化为对应的DFA。



步骤1：构造状态转化表

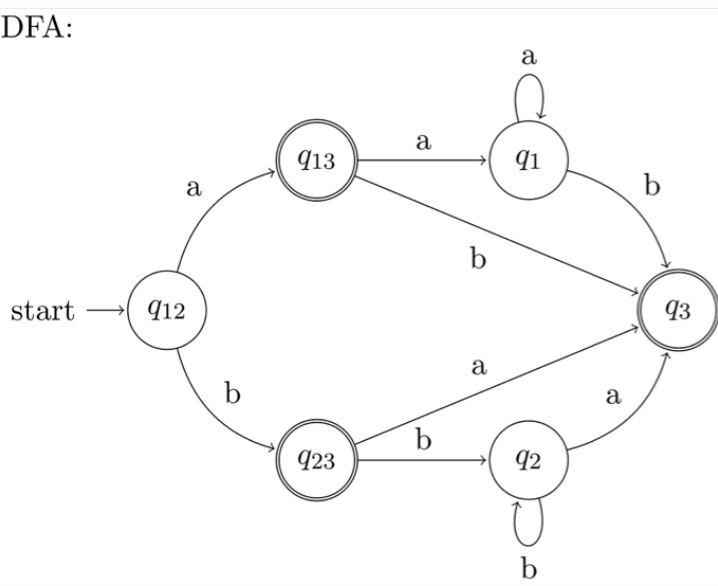
	a	b
状态0 = {q0 , q1, q2}	{q1, q3}	{q3, q2}
状态1 = {q1, q3}	{q1}	{q3}
状态2 = {q3, q2}	{q3}	{q2}
状态3 = {q1}	{q1}	{q3}
状态4 = {q3}	-	-
状态5 = {q2}	{q3}	{q2}

步骤2：确定DFA的起始状态和接收状态

起始状态： 状态0

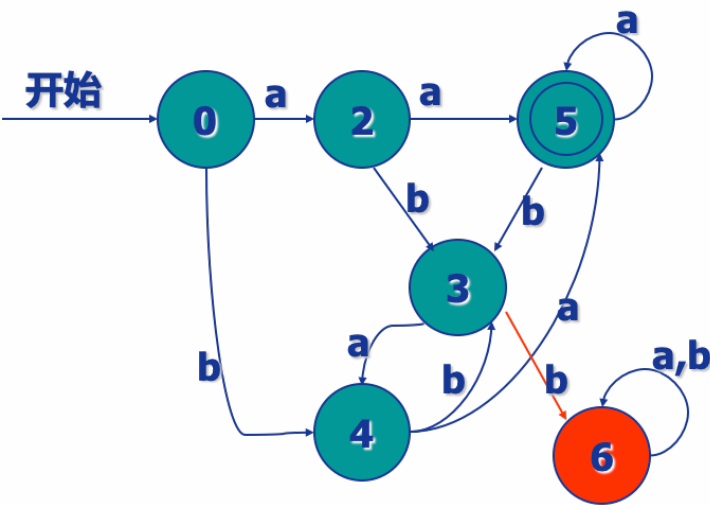
接收状态： 状态1、状态2、状态4

步骤3：根据状态转化表和起始状态、接收状态构造DFA



DFA化简

例题1：



步骤1：补充死状态

在进行化简前一定要确保所有状态都对于所有字母表中的每一个字母有一个转移，那些本身并不存在的转移都会汇聚到一个新的状态——死状态。

上图中的6状态就是一个死状态。

步骤2：把接收状态写到一个集合，非接受状态写到另一个集合

接收状态集合：{5}

非接受状态集合：{0, 2, 3, 4, 6}

请注意，死状态属于非接受状态。

步骤3：判断这两个集合中的元素是否具有不可区分性

	a	b
0	2	4
2	5	3
3	4	6
4	5	3
6	6	6

1. 从状态转化表中可以看到，状态2和状态4转换a后都进入了状态5，这是另一个集合，所以，状态2、4与其他元素有可区分性。

$\{2, 4\} \{0, 3, 6\} \{5\}$

2. 然后，我们发现状态2，4无论接收a还是接收b都会转化到相同的状态，所以状态2、4没有区分性，应该被化简合

并为一个状态。

3. 然后，我们看状态{0, 3, 6}中，{0,3}在接收a后会进入集合{2,4}，所以和{6}是可区分状态。

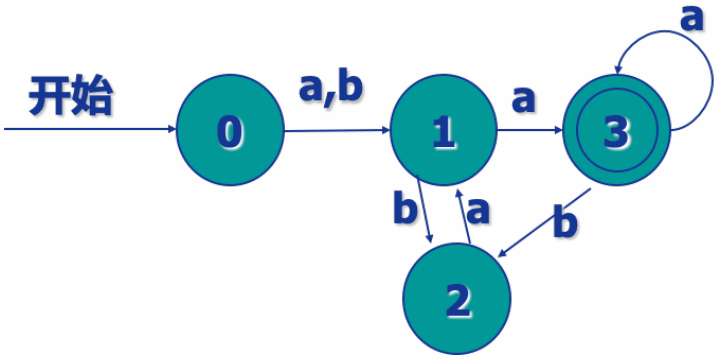
$$\{2, 4\} \{0, 3\} \{6\} \{5\}$$

4. 然后，我们看状态{0,3}中，{0,3}的状态表是不相同的，所以{0,3}也是可区分状态。

$$\{0\} \{3\} \{2, 4\} \{6\} \{5\}$$

我们就完成了化简，再感觉状态转换表画出化简后的DFA即可。

步骤4：构造DFA，删除死状态

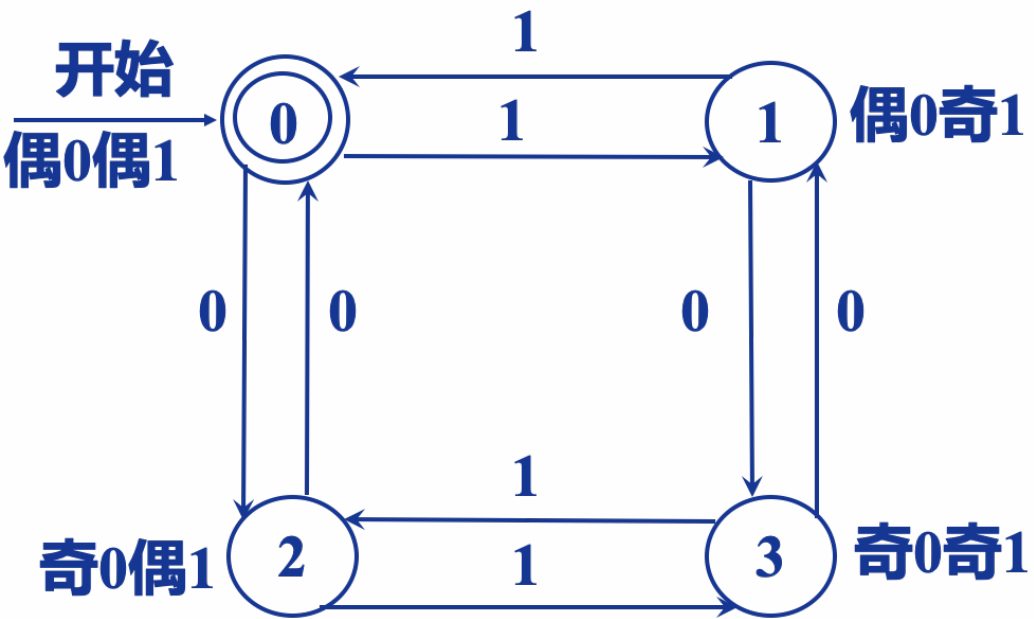


根据描述写出正则式

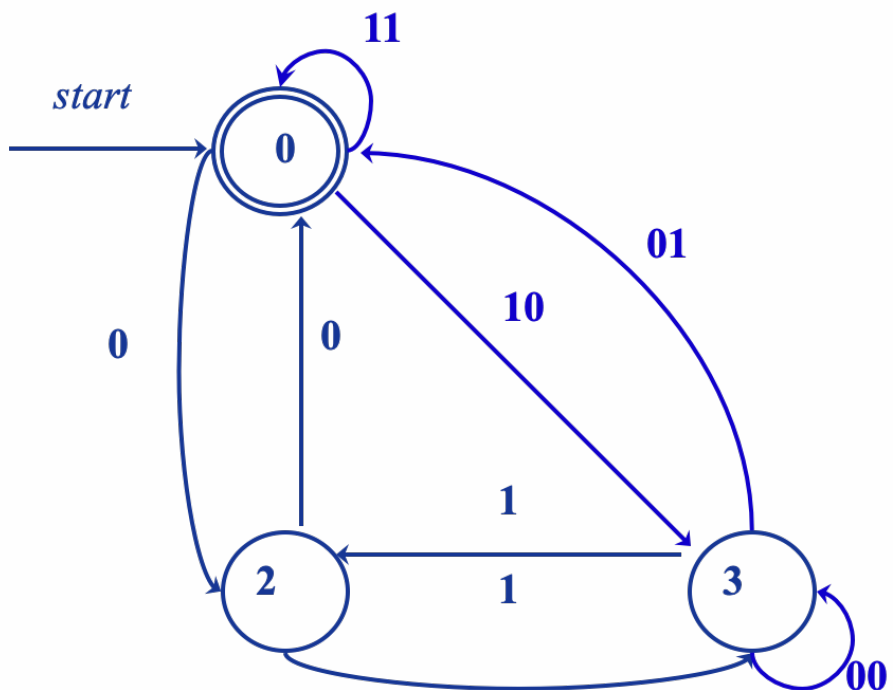
解法1：从DFA推导正规式

本方法适合那些状态个数少且状态易于划分的问题。

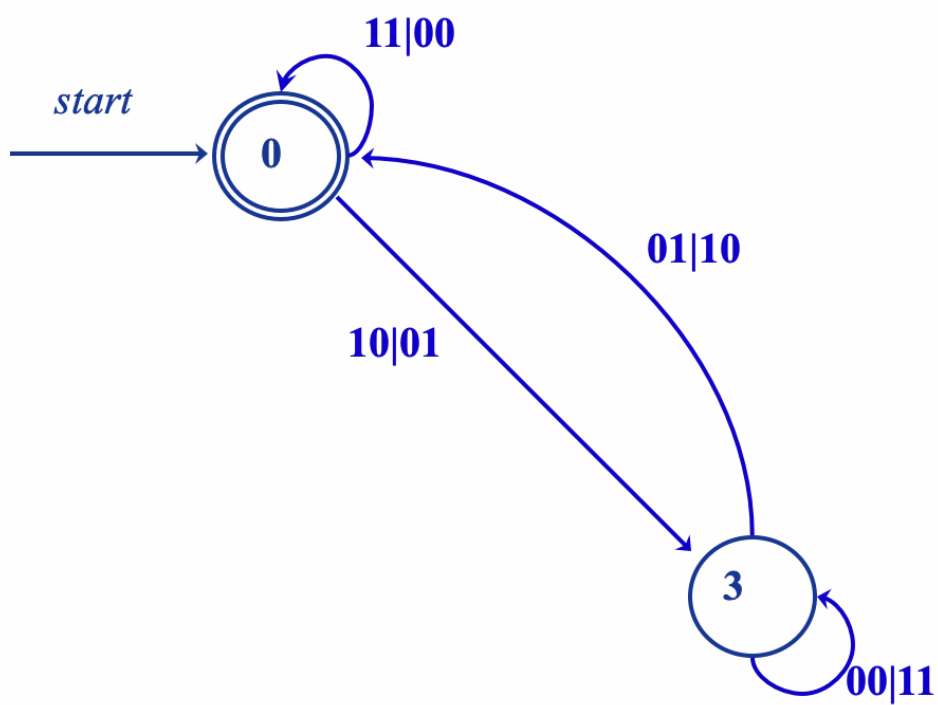
例1：接受 0和1的个数都是偶数的字符串



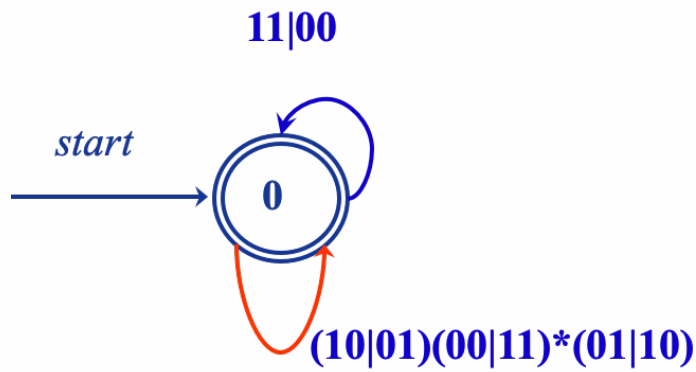
消除状态1：



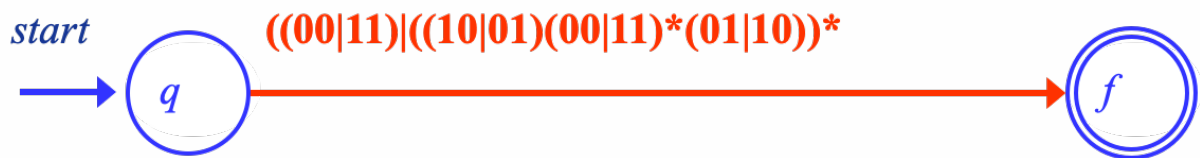
消除状态2:



消除状态3:



最终：



解法2：从小到大分析法

从逻辑上一点一点分析出来，先考虑符合条件的特例，再在特例上拼接直到涵盖所有情况为止。

例1：接受 0和1的个数都是偶数的字符串

1. 由若干个00或11开始；

$$(00|11)^*$$

2. 一旦出现一个01或10，那么经过若干个00或11后，一定会出现一个01或10。这第二个01或10的后面可能还有若干个00或11，一直到串的结束，或者到再次出现01或10为止。如果串没有结束的话，就是重复出现这里所描述的结构。

$$((01|10)(00|11)^*(01|10)(00|11)^*)^*$$

3. 所以正规式的最终结果由这两部分拼接而成：

$$(00|11)^*((01|10)(00|11)^*(01|10)(00|11)^*)^*$$

例2：不存在连续的两个0的二进制串：

1. 第一个字符如果是1，那么它后面可以跟若干个1（可以是0个1）直到出现0：

$$1^+0$$

2. 如果出现0，那么0后面必然是若干个1(至少是1个1)，直到再次出现0：

$$1^+0$$

3. 拼接上面的两种情况：

$$(1^+0)^*$$

4. 考虑最后一位，它可以是0，也可以是1。我们现在的表示只能是0，所以要考虑可以是1的情况：

$$((1^+0)^*1^+)^* = (1^+01^+)^*$$

5. 如果第一个字符是0，那么后面必然要接至少一个1，直到出现一个0为止：

$$(01^+)^*$$

5. 考虑最后一位还有可能是0的情况。请注意，如果是0，就只可能是最后一位的一个0：

$$(01^+)^*(0|1)$$

6. 合并4、6得到最终结果：

$$(1^+01^+)^*|(01^+)^*(0|1)$$

注：很有可能还有化简的余地，但是都是等价的。

第四大题

给出语法分析推导及语法分析树

答题指南

这部分大题非常的简单。但是一定要作答规范，有以下几点需要非常小心：

1. 一定看清题干是最左推导还是最右推导
2. 在推导时一定是优先推导最左/最右侧的非终结符，切勿随意！
3. 推导过程使用

$$a \Rightarrow_{lm} b \text{ or } a \Rightarrow_{rm} b$$

4. 在利用推导绘制语法树时，最左与最右一定要一以贯之。

例1：利用：

$$E \rightarrow E + E \mid E * E \mid (E) \mid - E \mid id$$

推导出

$$-(id + id)$$

解：

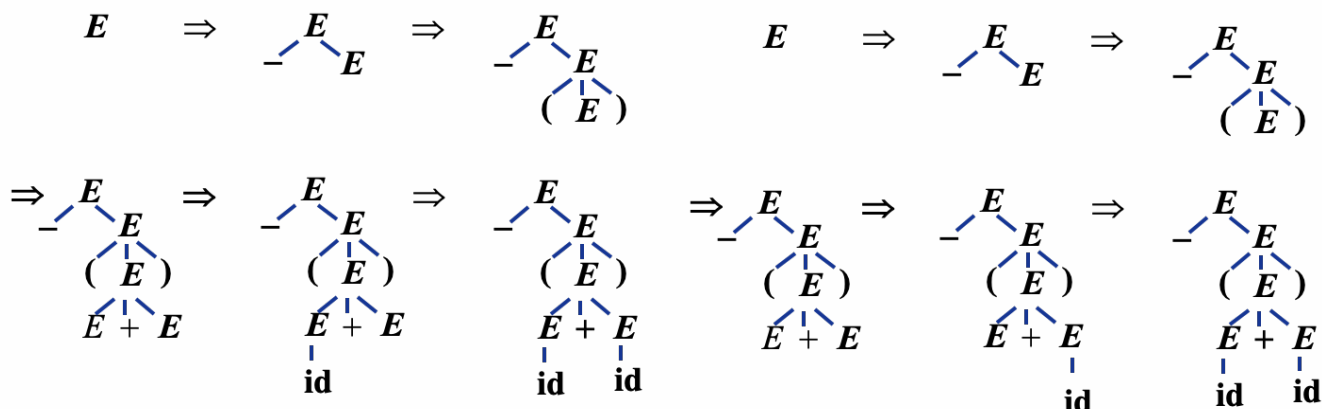
1. 最左推导：

$$\begin{aligned} E &\Rightarrow \text{Im } -E \Rightarrow \text{Im } -(E) \Rightarrow \text{Im } -(E + E) \\ &\Rightarrow \text{Im } -(\text{id} + E) \Rightarrow \text{Im } -(\text{id} + \text{id}) \end{aligned}$$

2. 最右推导:

$$\begin{aligned} E &\Rightarrow \text{Im } -E \Rightarrow \text{Im } -(E) \Rightarrow \text{Im } -(E + E) \\ &\Rightarrow \text{Im } -(\text{id} + E) \Rightarrow \text{Im } -(\text{id} + \text{id}) \end{aligned}$$

3. 语法树



我们观察到，如果文法不存在二义性，那么最左推导与最右推导的结果应该完全一样，可以利用这种性质进行验算。

例2 【2018年真题】：

$$\begin{array}{l} E \rightarrow T \mid E + T \\ T \rightarrow F \mid T * F \\ F \rightarrow (E) \mid i \end{array}$$

1. 请写出 (T^*F+i) 的最右推导

$$E \Rightarrow_{rm} T \Rightarrow_{rm} F \Rightarrow_{rm} (E) \Rightarrow_{rm} (E + T) \Rightarrow_{rm} (E + F) \Rightarrow_{rm} (E + i) \Rightarrow_{rm} (T + i) \Rightarrow_{rm} (T * F + i)$$