

期中测验参考答案

一、判断题（10 分）

1. (F) 语言 $\{a^n b^m \mid n \neq m\}$ 可以用正则表达式描述。
2. (T) $LL(k)$ 对应最左推导， $LR(k)$ 对应最右推导。
3. (T) 如果一个文法的某一个字符串有两种不同的解析树，那么该文法是二义的。
4. (T) $LL(1)$ 语法分析不能处理带有左递归的上下文无关文法。
5. (F) Burke-Fisher 方法是一种使用额外的栈和队列来保存并恢复到旧状态的局部错误恢复方式。
6. (F) 语言 $\{s \in \{a, b\}^* \mid s \text{ 含有相同数量的 } a \text{ 和 } b\}$ 可以使用文法 $S \rightarrow aSb \mid bSa \mid \varepsilon$ 描述。
7. (T) 对于产生式 $A \rightarrow B\alpha C$ ，有 $\text{Follow}(A) \subseteq \text{Follow}(C)$ 和 $\text{First}(B) \subseteq \text{First}(A)$ 。
8. (F) 在构造 $LR(0)$ 分析表时，如果一格内出现两条规约动作，我们可以通过消除左递归的方式改进语法来处理冲突。
9. (F) 当 k 足够大时， $LL(k)$ 可以表示所有 $LR(0)$ 文法能表示的语言。
10. (F) 在命令式符号表更新中， $\{a \mapsto \text{int}\} + \{a \mapsto \text{string}\} = \{a \mapsto \text{int}, a \mapsto \text{string}\}$ 。

解答:

1. 题中正则表达式表示的语言等价于 $\{a^* b^*\} - \{a^n b^n\}$ ，其中 $\{a^n b^n\}$ 是 CFG。
2. 可能比较容易错的地方是 $LR(k)$ ，对应的是一个逆向的最右推到过程。
3. 文法非歧义的定义是，对于每一个字符串，存在唯一的解析树；或者等价的，对于每一个字符串，存在唯一的最左推导/最右推导。
4. 略
5. 略
6. 根据 First 和 Follow 的定义。

7. 字符串 $abba$ 不能被该文法识别

8. 消除左递归是用于处理 LL(1) 语法中存在的问题的方法。

9. LR(1) 就存在一些子集不能被 LL(k) 识别，顺便在第三章课后练习作业中，有一个 $L \in \text{LL}(1)$ 但是 $L \notin \text{LALR}(1)$ 的例子。

10. 在命令式符号表更新时如果更改了原有符号的类型，则更新后原来的符号类型不可见。

二、单项选择题 (20 分)

11. (C) 下列哪个环节不属于典型“前-中-后”三段式编译器中的前端？

A. 词法分析 B. 语法分析 C. 指令选择 D. 语义分析

12. (A) 下列哪个文法是自顶向下的？

A. LL(1) B. LR(0) C. SLR(1) D. LALR(1)

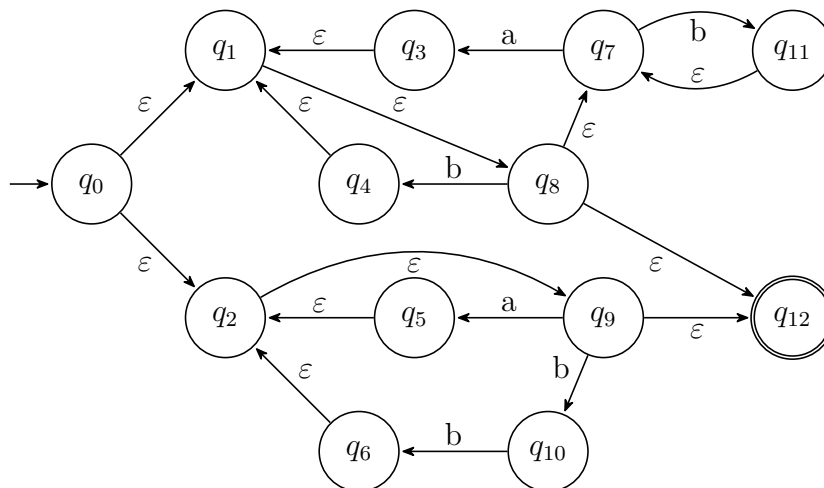
13. (A) 默认情况下，Yacc/Bison 在遇到移入-归约冲突时，采取的策略是

A. 移入 B. 归约 C. 报告错误 D. 使用在文法描述中先出现的产生式 D

14. (D) 给定上下文无关文法 $S \rightarrow B\$$, $B \rightarrow B(B) \mid \varepsilon$, 则 $\text{Follow}(B)$ 为

A. $\{\$ \}$ B. $\{ (, \$ \}$ C. $\{ (, \$, \varepsilon \}$ D. $\{ (,), \$ \}$

15. (A,C) 给出如下 NFA，下列哪个正则表达式表示的语言可被该 NFA 接受？



A. $(b^*a \mid b)^* \mid (bb \mid a)^*$

B. $((ba)^* \mid b)^* \mid (bb \mid a)^*$

C. $(b^*a \mid b)^* \mid (bb \mid a^*)^*$

D. $((ba)^* \mid b)^* \mid (bb \mid a^*)^*$

16. (B) 下列哪个文法是 LL(1) 文法?

A. $S \rightarrow a \mid b \mid (T)$

$T \rightarrow T, S \mid S$

B. $A \rightarrow a N$

$N \rightarrow A B e \mid \varepsilon$

$B \rightarrow d M$

$M \rightarrow b M \mid \varepsilon$

C. $A \rightarrow b a B \mid \varepsilon$

$B \rightarrow a a A b \mid a a$

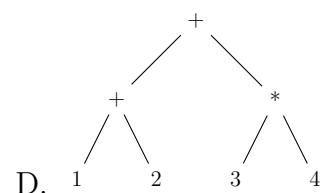
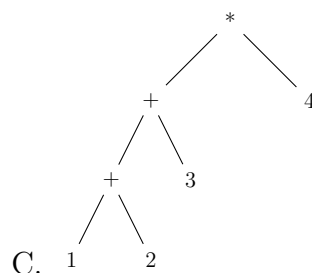
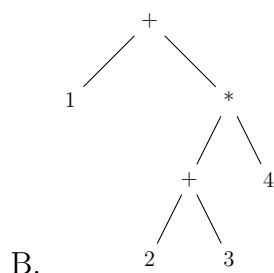
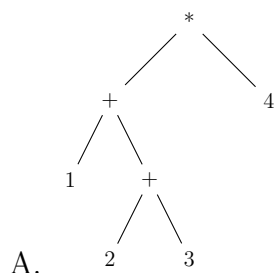
D. $S' \rightarrow S V$

$S \rightarrow a \mid b$

$a V \rightarrow a e$

$b V \rightarrow b f \mid b$

17. (B) 对于 C 语言风格的表达式语法和优先级, 下列哪个最可能是算术表达式 $1 + (2 + 3) * 4$ 的抽象语法树?



18. (A) 给定文法 $S \rightarrow A B \mid b C$, $A \rightarrow b \mid \varepsilon$, $B \rightarrow a D \mid \varepsilon$, $C \rightarrow A D \mid b$, $D \rightarrow a S \mid c$, 下面正确的是:

A. $\text{First}(A) = \{b\}$, $\text{Follow}(A) = \{a, c\}$

B. $\text{First}(B) = \{a, c\}$, $\text{Follow}(B) = \emptyset$

C. $\text{First}(C) = \{b, c\}$, $\text{Follow}(C) = \{a, b, c\}$

D. $\text{First}(D) = \{a, c\}$, $\text{Follow}(D) = \{a, c\}$

19. (D) 给定文法 $S \rightarrow b N$, $N \rightarrow B a N \mid \varepsilon$, $B \rightarrow a b$ 及其 LL(1) 分析表:

考虑输入串 $babaaba\$$, 下列哪个分析状态可能出现在 LL(1) 语法分析栈中 (栈底在右侧)?

A. $B a N N \$$

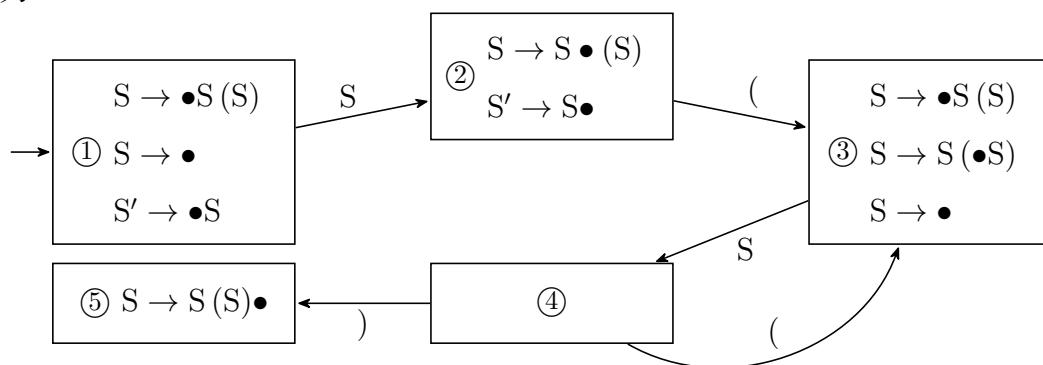
B. $B a B a \$$

	a	b	\$
S		$S \rightarrow bN$	
B	$B \rightarrow a b$		
N	$N \rightarrow B a N$		$N \rightarrow \epsilon$

C. Nb\$

D. BaN\$

20. (B) 给定文法 $S' \rightarrow S, S \rightarrow S(S) \mid \epsilon$, 考虑 SLR(1) 语法分析的自动机, 状态 ④ 应该为



- A. $S \rightarrow S \bullet (S)$ B. $S \rightarrow S \bullet (S)$
 $S \rightarrow S(S \bullet)$ C. $S \rightarrow S(S \bullet)$ D. $S \rightarrow S(S \bullet)$
 $S' \rightarrow \bullet S$

三、正则表达式、NFA 和 DFA (20 分)

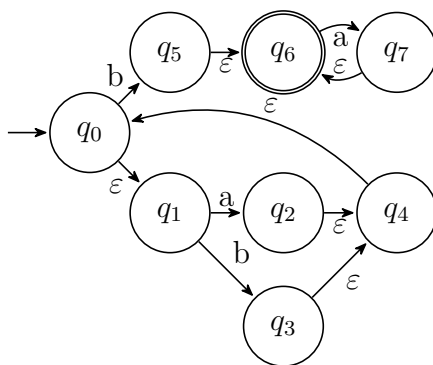
给定如下正则表达式：

$$(a|b)^*ba^*$$

1. 采用 Thompson 方法构造该正则表达式对应的 NFA。
2. 采用子集构造法将上述 NFA 转化为等价的 DFA 并最小化。

解答：

1. NFA 如下

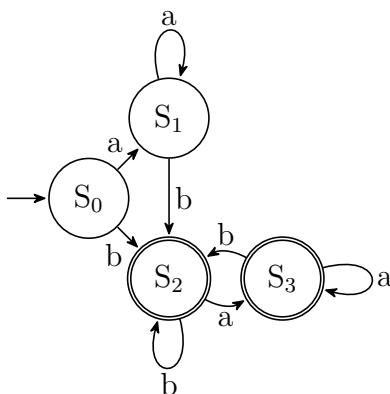


2. 子集构造过程

State	Closure
S ₀	{q ₀ , q ₁ }
S ₁	{q ₀ , q ₁ , q ₂ , q ₄ }
S ₂	{q ₀ , q ₁ , q ₃ , q ₄ , q ₅ , q ₆ }
S ₃	{q ₀ , q ₁ , q ₂ , q ₄ , q ₆ , q ₇ }

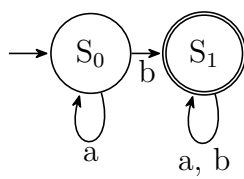
State	Transformation	
S ₀	a	S ₁
	b	S ₂
S ₁	a	S ₁
	b	S ₂
S ₂	a	S ₃
	b	S ₂
S ₃	a	S ₃
	b	S ₂

构造得到的 DFA：



其中 S_0 和 S_1 等价, S_2 和 S_3 等价。

化简后最终结果为:



四、LR(1) 语法分析 (40 分)

给定以下文法：

$$S' \rightarrow S \$$$

$$S \rightarrow A a$$

$$A \rightarrow B C \mid B C f$$

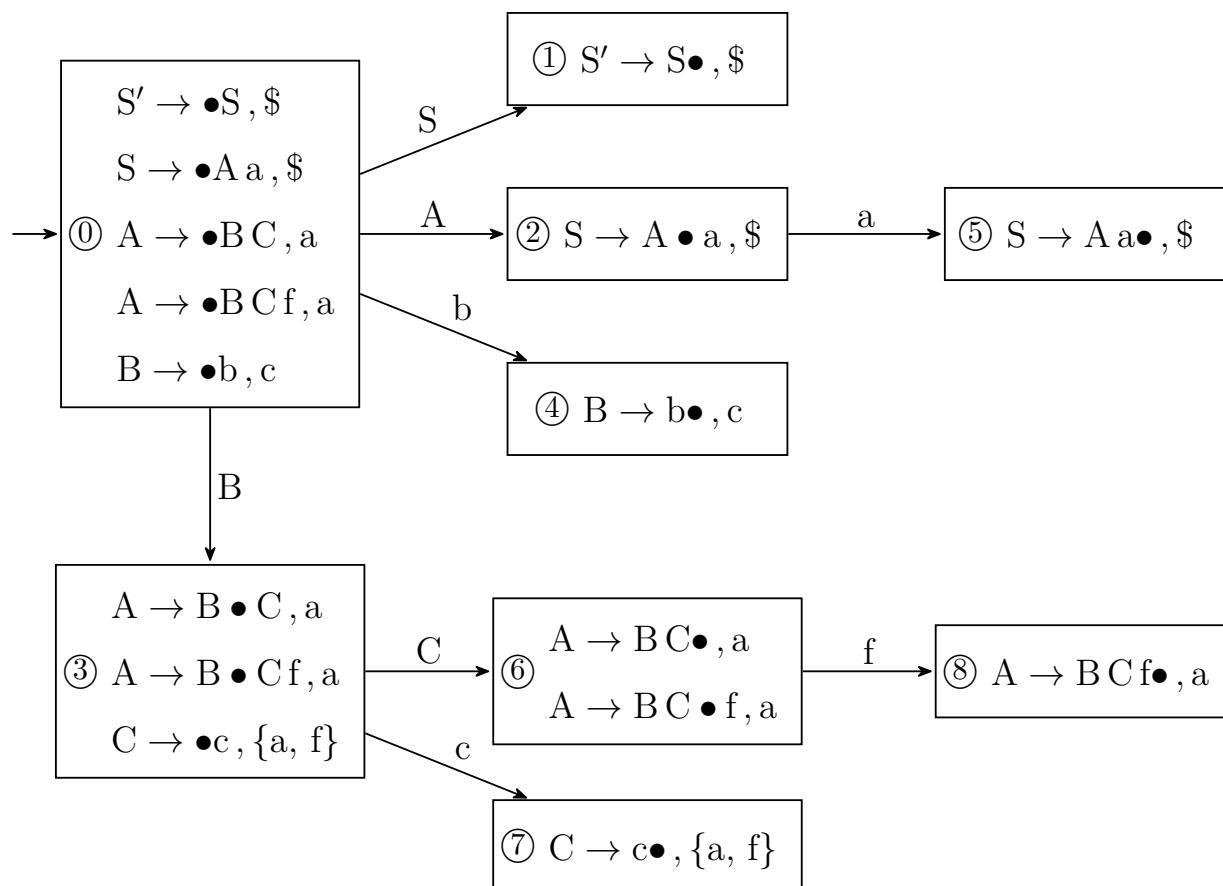
$$B \rightarrow b$$

$$C \rightarrow c$$

1. 构造该文法的 LR(1) 项并给出自动机。
2. 给出该文法的 LR(1) 语法分析表。
3. 判断该文法是否为 LR(1) 文法并说明理由。

解答：

自动机：



解析表：

	Action					Goto			
	\$	a	f	b	c	S	A	B	C
S_0				s4		g1	g2	g3	
S_1	acc								
S_2		s5							
S_3					s7				g6
S_4					r5				
S_5	r2								
S_6		r3	s8						
S_7		r6	r6						
S_8		r4							

因此，这是 LR(1) 文法