清华大学本科生考试试题专用纸

考试课程 《形式语言与自动机》 A 卷 2019 年 6 月 23 日

学号：______________ 姓名：______________ 班级：______

（注：解答可以写在答题纸上，也可以写在试卷上；交卷时二者都需要交回。）

一. (16 分) 判别下列各命题的真假性，回答 true 或者 false: (每小题 2 分)

1．存在算法可以判定“两个正规语言是否相交”。

____

2．针对正规语言的 Pumping 引理可用来判定某一语言是正规语言。

____

3．如果$L_1 \subseteq L_2$ 且 L_2 是正规语言，那么 $L_2 - L_1$ 也是正规语言。

____

4．存在一个算法，可用于判定一个正规表达式的语言是否为无穷语言。

____

5．非固有二义的上下文无关语言，一定是某个 *DPDA* 的语言。

____

6．设问题*A*可多项式时间归约到问题*B*。若*B*是*NP*-完全问题，则*A*也是*NP*-完全问题。

____

7．可以用带两个计数器的计数器机模拟任何一个带两个以上下推栈的多栈机。

____

8．通用语言所对应的问题可归约到以下问题：“任意给定的 CFG 是否二义文法？”。

____

二. (12 分) 选择填空 (每小题 2 分)

1．下列语言中，__________________ 是正规语言。

供选择的答案：

A. $\{ xwx^R \mid x, w \in \{ a, b \}^+,\ x^R$ 为 x 的反向 $\}$

B. $\{ xcx^R \mid x \in \{ a, b \}^+,\ x^R$ 为 x 的反向 $\}$

C. $\{ xx^R \mid x \in \{ a, b \}^+,\ x^R$ 为 x 的反向 $\}$

D. 以上都不是。

2. 下列语言中，________________ 是某个 DPDA 的语言，但不是某个空栈接受的 DPDA 的语言，也不是正规语言。

供选择的答案：

A. $\{\, xwx \mid x, w \in \{\, a, b\,\}^{+}\,\}$

B. $\{\, xcx \mid x \in \{\, a, b\,\}^{+}\,\}$

C. $\{\, xxw \mid x, w \in \{\, a, b\,\}^{+}\,\}$。

D. 以上都不满足。

3. 下列语言中，________________ 是某个 PDA 的语言，但不是某个 DPDA 的语言。

供选择的答案：

A. $\{\, xwx^R \mid x, w \in \{\, a, b\,\}^{+}$, x^R 为 x 的反向 }

B. $\{\, xcx^R \mid x \in \{\, a, b\,\}^{+}$, x^R 为 x 的反向 }

C. $\{\, xx^R \mid x \in \{\, a, b\,\}^{+}$, x^R 为 x 的反向 }

D. 以上都不是。

4. 下列语言中，________________ 是某个可停机图灵机的语言，但不是某个 PDA 的语言。

供选择的答案：

A. $\{\, w \mid w\in\{\, a, b\,\}^{+}$，且 w 不为 xx 的形式 }。

B. $\{\, xx \mid x\in\{\, a, b\,\}^{+}\,\}$。

C. $\{\, w \mid w\in\{\, a, b\,\}^{+}$，且 x 不为 xx^R 的形式，其中 x^R 为 x 的反向 }。

D. 以上都不是。

5. 下列语言中，________________ 是某个图灵机的语言，但不是某个可停机图灵机的语言。

供选择的答案：

A. 语言 L_d（课程中定义的对角语言）

B. 语言 L_u（课程中定义的通用语言）

C. $\{\, xx^Rx^Rx \mid x \in \{\, a, b\,\}^{+}$, x^R 为 x 的反向 }

D. 以上都不是。

6. 下列语言中，________________ 是某个图灵机的语言。

A. 语言 L_u（课程中定义的通用语言）的补语言

B. 语言 L_H（课程中图灵机停机问题所定义的语言）的补语言

C. 语言 L_H（课程中图灵机停机问题所定义的语言）

D. 以上都不是。

三. (32 分) 简答题:

1. (4 分) 设 CFG G =（{S, A, B, C, D}, { a, b, c }, P, S），其中 P 由下列产生式构成：

$$\begin{aligned} &S \to AB \mid \varepsilon \\ &A \to CA \mid a \\ &B \to Bb \mid \varepsilon \\ &C \to B \\ &D \to AC \mid c \end{aligned}$$

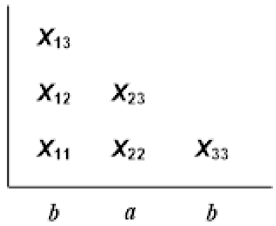
（1）消去P中的ε-产生式得到产生式集合P_1，构成 CFG G'，使得 $L(G') = L(G) - \{\varepsilon\}$. 给出$P_1$ = ？ （2分）

（2）消去P_1 中的Unit 产生式得到产生式集合P_2，构成 CFG G''，使得$L(G'') = L(G')$. 给出P_2 = ？ （1分）

（3）消去 P_2 中的无用符号得到产生式集合P_3. 给出P_3 = ？ （1分）

2. (4 分) 文法 G（S为开始符号）的产生式集合为：

$$\begin{aligned} &S \to SB \mid AB \\ &A \to BA \mid a \mid b \\ &B \to AA \mid b \end{aligned}$$



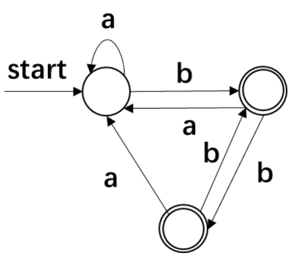
上图表示对于文法 G 和字符串 bab 应用 CYK 算法时所构造的表。

(1) 分别计算图中所有 X_{ij} （$1 \le i, j \le 3$）

(2) 是否有 $bab \in L(G)$?

3. (6分) 下图表示一个有限自动机 A：

（注意：本题中多处出现有限自动机的描述，可以是也可以不是 DFA）



(1) 试采用课程中所介绍的方法，给出一个有限自动机 B，使得 $L(B) = (L(A))^R$（2分）

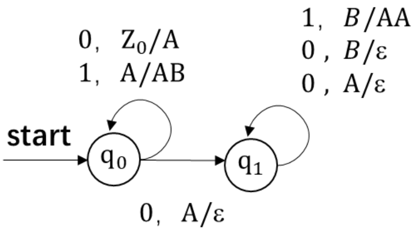
（$(L(A))^R$ 为 $L(A)$ 的反向）

(2) 试采用课程中的方法，给出一个有限自动机 C，使得 $L(C) = \{a,b\}^* - L(A)$ （2分）

(3) 设映射 h：{ 0,1 } $\to \{a,b\}^*$ 定义为 $h(0) = ab$, $h(1) = ba$；

试构造一个有限自动机 D，使得 $L(D) = h^{-1}(L(A))$。 （2分）

4. (3分) 下图刻画了 PDA $P = (\{q_0, q_1\}, \{0,1\}, \{Z_0,A,B\}, \delta, q_0, Z_0)$ 的转移规则，试严格利用课程中介绍的从空栈接受的 PDA 到 CFG 的转换算法，定义一个与该 PDA 等价的 CFG，开始符号设为 S：



5. (2 分) 已知若 L 和 M 为正规语言，则L和M的并 L∪M也是正规语言，且已知若L为正规语言，则L的补 $\overline{L} = \Sigma^* - L$ 也是正规语言。仅用这两个条件，证明若 L 和 M 为正规语言，则 L和M的差L－M 也是正规语言.

6. (4 分) 将下面的上下文无关文法转换为等价的以空栈方式接受的PDA：

$$S \rightarrow ABa \mid BAb \mid a$$
$$A \rightarrow Aa \mid a$$
$$B \rightarrow Sb \mid b$$

7. (5分) 对于语言

$L = \{ w \mid w \in \{a,b,c\}^*$，且在$w$中，$a$的数目等于$b$的数目，c的数目大于$a$的数目$\}$

可以利用 Pumping 引理证明 L 不是上下文无关语言，以下是一个证明概要：

考虑任意的 $n \geq 1$。取 $z =$ ___①___ $\in L$ 。
对任意满足条件 $z = uvwxy \wedge vx \neq \varepsilon \wedge |vwx| \leq n$ 的 u, v, w, x, y,
若 _②_ 时，取 $k =$ _③_；
若 _④_ 时，取 $k =$ _⑤_。（若需更多分支，可自行添加）
则有 $uv^kwx^ky \notin L$。

试在其中 ①、②、③、④ 和 ⑤ 处填写适当的内容。（若需更多分支，可自行添加）

8. (4 分) 证明$\{0^n \mid n = m^2, m$是正整数$\}$不是正规语言。

四. (25 分) 设计题：（必要时解释设计思路）

1. (5 分) 试构造接受下列语言的一个有限自动机（DFA，NFA，ε－NFA均可），要求状态数不超过 7：

$L = \{ w \mid w \in \{a, b\}^*$，$|w|>=3$ 并且 w 的最后 3 个字符中包含偶数个 b $\}$

注：要求状态数不超过 7，并不意味着状态数一定会达到 7。后面的题目也类似。

2. (5 分) 试给出下列正规语言的一个正规表达式：

$L = \{ w \mid w \in \{0,1\}^*$，$w$至多含有2个子串01$\}$

3. (5 分) 试给出下列语言的一个上下文无关文法：

$L = \{ a^nb^m \mid 3m \leq 2n \leq 5m \}$

4. (5 分) 试构造接受下列语言的一个 PDA（终态接受和空栈接受均可，必要时给出设计思路），要求该 PDA 的堆栈符号数不超过 3，且用状态转移图描述你的设计：

$L = \{ w \mid w = a^i b^j c^k$，其中 $i \neq 2j$ 或 $j \neq 2k \}$

5. (5 分) 试构造一个可停机图灵机 $M = (Q, \{0, 1\}, \{0, 1, B\}, \delta, q_0, B, \{q_f\})$。以任意长度 $n > 0$ 的 0、1 串为输入，图灵机 M 执行后（到达终态）的效果为：（1）带头回到原来首字符的位置（起始单元格）；（2）该串的首字符被丢弃，剩余字符串整体向左移动一个单元格，即原第2个字符移至原第1个字符的单元格，原第3个字符移至原第2个字符的单元格，…，原第 n 个字符移至原第 n-1 个字符的单元格；（3）原第 n 个字符的单元格填入字符 0，即末位补0，使带上的（输出）结果0、1 串长度维持原输入串长度（仍为n）。要求该图灵机 M 从开始执行至停机的移动步数与输入串长度 n 成线性关系，即时间复杂度为O(n)。用状态转移图描述你所设计的图灵机。

五. (15 分) 证明题: （要求证明过程严谨，步骤明确。）

1. (5 分) 已知语言 $N = \{ a^n \mid n = i*i, i \geq 0\}$ 不是CFL，试利用该结论以及CFL的封闭运算，证明如下语言 L 不是CFL：

$L = \{ w \mid w \in \{a,b,c\}^*$，$w$ 中 a, b 和 c 的个数分别为 $\#a, \#b$ 和 $\#c$，满足$\#b = (\#a + \#c)^2 \}$

2. (5 分) 试证明：如果语言 L 是某个空栈接受的 DPDA P 的语言 $N(P)$，那么 L 一定满足前缀性质，并且 L 是某个 DPDA P' 的语言 $L(P')$。

3. (5 分) 设 DFA $A = (Q, \Sigma, \delta, q_0, F)$，对任何 $p, q \in Q$,，称 p 和 q 是不可区别的，当且仅当 $\forall w \in \Sigma^*. (\hat{\delta}(p,w) \in F \leftrightarrow \hat{\delta}(q,w) \in F)$；否则就是可区别的。课程中介绍了一种填表算法，可以基于如下归纳过程标记出全部可区别的状态偶对：

（1）基础。如果 p 为终态，而 q 为非终态，则 p 和 q 标记为可区别的。

（2）归纳。设 p 和 q 已标记为可区别的，如果状态 r 和 s 通过某个输入符号 a 可分别转移到 p 和 q，即有 $\delta(r,a)=p$ 和 $\delta(s,a)=q$，则 r 和 s 也标记为可区别的。

试证明：若 Q 中的两个状态是可区别，则二者一定会被上述算法标记为可区别的。

提示：以下列举课程中叙述过的有关可区别性的几个结论，需要时可以直接应用：

设状态 r 和 s 通过某个输入符号 a 可分别转移到 p 和 q（即 $\delta(r,a)=p, \delta(s,a)=q$），则有：

（1）p 和 q 可区别 $\Rightarrow r$ 和 s 可区别；

（2）r 和 s 不可区别 $\Rightarrow p$ 和 q 不可区别；

（3）r 和 s 可由 ax 区别 $\Rightarrow p$ 和 q 可由 x 区别。