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 2013 年 6 月 22 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. 若 $L_1 \cap L_2$ 是正规语言，则 L_1 和 L_2 一定都是正规语言。

____

2. 若 L_1 和 L_2 都是正规语言，则 $L_1 \cap L_2$ 一定是正规语言。

____

3. 若 $L_1 \cup L_2$ 不是正规语言，则 L_1 和 L_2 都不是正规语言。

____

4. 判定一个串不能被某个有限自动机接受的算法是存在的。

____

5. 存在一个判定任意两个正规表达式是否拥有相同语言的算法。

____

6. 一个递归可枚举语言和它的补语言不可能都是递归语言。

____

7. 非确定图灵机的语言所对应的问题是一个NP问题。

____

8. 任何多带（multi-tape）图灵机均可由一个多道（multi-track）图灵机来模拟。

____

二. (12 分) 选择填空 (每小题 2 分)

1. 语言 $\{0^n 2 1^m \mid n \geq m\}$ ________________。

2. 语言 $\{0^n 2 1^m \mid n \geq 1, m \geq 1，\mathrm{n} + \mathrm{m} \leq 100\}$ ________________。

3. 语言 $\{0^n 1^m 2 \mid n \geq 1, m \geq 1，\mathrm{n} + \mathrm{m} \leq 100\}$ ________________。

4. 语言 $\{ww^R \mid w \in \{0, 1\}^*$，$w^R$ 为 w 的反向 $\}$ ________________。

5. 语言 $\{w2w \mid w \in \{0, 1\}^*\}$ ________________。

6. 语言 $\{w2w^R \mid w \in \{0, 1\}^*$，$w^R$ 为 w 的反向$\}$ ________________。

供 1、2 和 3 三小题选择的答案：

A. 是某个有限自动机的语言，也是某个空栈接受方式的 *DPDA* 的语言.

B. 是某个有限自动机的语言，但不是任何空栈接受方式的 DPDA 的语言。

C. 既是某个终态接受方式的 *DPDA* 的语言，又是某个空栈接受方式的 *DPDA* 的语言，但不是任何有限自动机的语言。

D. 是某个终态接受方式的 DPDA 的语言, 但不是任何空栈接受方式的 DPDA 的语言, 也不是任何有限自动机的语言。

E. 是某个 PDA 的语言，但不是任何 DPDA 的语言。

F. 不是任何 PDA 的语言。

三. (32 分) 简答题:

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

$$S \to ABC \mid \varepsilon$$
$$A \to BS \mid a$$
$$B \to b$$
$$C \to \varepsilon$$

（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：

$$S \to AB \mid BC$$
$$A \to BA \mid a$$
$$B \to CC \mid b$$
$$C \to AB \mid a$$

下图表示对于文法 G 和字符串 *abba* 应用 CYK 算法时所构造的表（部分X_{ij}已给）。

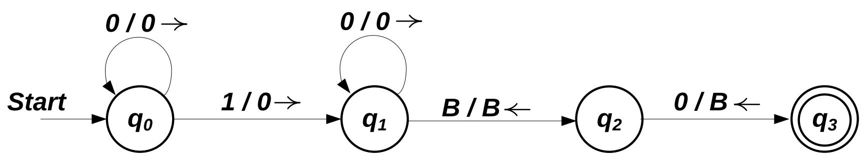
(1) 分别计算出 X_{14} 和 X_{22};　（3分）

(2) 是否有 $abba \in L(G)$?　（1分）

X_{14}			
X_{13}	X_{24}		
X_{12}	X_{23}	X_{34}	
X_{11}	X_{22}	X_{33}	X_{44}
a	***b***	***b***	***a***

X_{14}			
{ }	{A}		
{S,C}	X_{23}	{S,A}	
{A,C}	X_{22}	X_{33}	{A,C}
a	***b***	***b***	***a***

3. (4分) 下图描述了图灵机 $M = (\{q_0, q_1, q_2, q_3\}, \{0, 1\}, \{0, 1, B\}, \delta, q_0, B, \{q_3\})$:



(1) 指出该图灵机的语言 $L(M)$　　　　（2分）

(2) 指出对于任何 $w\in L(M)$，该图灵机到达终态时带上的内容。　　　　（2分）

4. (4 分) 设有空栈接受方式的 $PDA\ P = (Q, S, \Gamma, \delta, q_0, Z_0)$。可以定义一个等价于$P$的终态接受方式的 $PDA\ P' = (Q \cup \{p_0, p_f\}, S, \Gamma \cup \{X_0\}, \delta', p_0, X_0, \{p_f\})$，即定义满足$L(P') = N(P)$ 的$PDA\ P'$。其中，

$\delta'(p_0，\varepsilon，X_0) =$ ___①___ ，

以及对任何 $q\in Q$，

$\delta'(q，\varepsilon，X_0) =$ ___②___ 。

5. (4 分) 以下 2 组产生式分别对应 2 个文法 G 和 G_1 的定义（开始符号均为 S）：

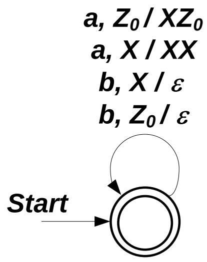
G： $S \rightarrow aAA \mid \varepsilon$
$A \rightarrow aS \mid bS \mid a$

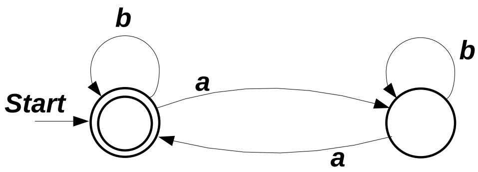
G_1：$S \rightarrow 0S1 \mid \varepsilon$

并设有 $\{0, 1\}$ 上的语言 $M = \{01, \varepsilon\}$

设替换映射 s：$\{a, b\} \rightarrow \{0, 1\}$ 定义为：$s(a) = L(G_1)$，$s(b) = M$。试给出语言 $s(L(G))$ 的一个上下文无关文法。

6. (4 分) 下面左图描述一个 $PDA\ P$，右图描述一个 $DFA\ A$：





试构造一个 $\{a, b\}$ 上的 $PDA\ P'$，使得 $L(P') = L(P) - L(A)$。

7. (4 分) 对于 $\{a, b\}$ 上的语言 $L = \{ w \mid w \in \{a, b\}^*, w$ 的任何后缀中 a 的个数不超过 b 的个数 $\}$，以下是利用 Pumping 引理证明 L 不是正规语言的一个证明概要：

考虑任意的 $n\geq 1$。取 $w =$ ___①___ $\in L$。
对任意满足条件 $w=xyz \wedge y\neq\varepsilon \wedge |xy|\leq n$ 的 x, y, z,
若取 $k =$ _②_，则有 $xy^kz \notin L$。

试在其中 ① 和 ② 处填写适当的内容。

8. (4 分) 对于语言 $L = \{ w \mid w \in \{a, b, c, d\}^*, w$ 中 a 的个数等于 b 的个数且 c 的个数等于 d 的个数 $\}$，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,
取 $k =$ _②_，则有 $uv^kwx^ky \notin L$。

试在其中 ① 和 ② 处填写适当的内容。

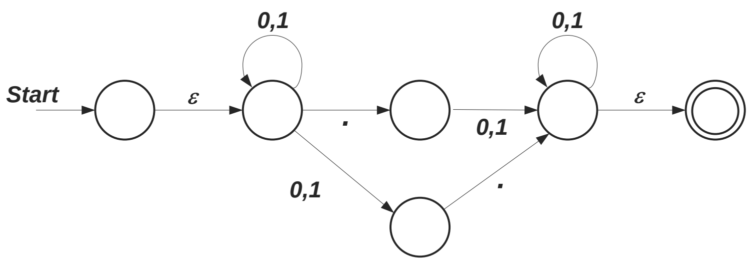
四. (25 分) 设计题: （要求适当解释设计思路）

1. (5 分) 试构造接受下列语言的一个确定有限自动机（DFA），且该有限自动机的状态数不超过 5：

$$L = \{ w \mid w \in \{a, b\}^*，w \text{ 中 } a \text{ 的个数是偶数，且 } w \text{ 的长度也为偶数} \}$$

注：要求状态数不超过 5，并不意味着状态数一定会达到 5。后面的题目也类似。

2. (5 分) 下图中的 ε-NFA 描述了字母表{ 0，1，**.** }上的正规语言，用于表示某种合法的二进制小数集合。试给出该语言的一个正规表达式，且该表达式中运算符的总数不超过 20（只能使用 '+', '*' 以及 '连接' 3 种运算符和括号，不计括号数）。



3. (5 分) 试给出下列语言的一个上下文无关文法，且该文法的非终结符数目不超过 8：

$$L = \{a^n b^i c^j d^m \mid n, m, i, j \geq 0 \wedge n + m = i + j\}$$

4. (5 分) 试构造接受下列语言的一个 PDA（空栈接受或终态接受均可），要求该 PDA 的状态数和堆栈符号数均不超过 5，并且每一步转移中栈顶符号最多可替换为两个符号：

$$L = \{ w \mid w \in \{a, b, c\}^*，w \text{ 中 } a \text{ 和 } b \text{ 的个数相同且不含连续的 } c \}$$

5. (5 分) 试设计一个图灵机 $M = (Q, \{ 0,1\}, \{ 0, 1, B \}, \delta, q_0 ,B , \{ q_f\})$ 可以将二进制形式的非负整数 n 作为输入，并作如下计算：若 n 为偶数，则输出结果为 $n+1$；若 n 为奇数，则输出结果为 $2n$ 。开始时 M 处于状态 q_0，带中包含着二进制数 n，其它单元格均为 B，带头正扫描 n 的最左一位。所设计的图灵机 M 应当停机。停机时，M 处于状态 q_f，带上即为上述计算结果的二进制形式，而其它单元格均为 B。到达状态 q_f 时，带头处于何处不作要求。用状态转移图描述你所设计的图灵机。

五. (15 分) 证明题:

1. (5 分) 已知语言 $L_{01} = \{ 0^k1^k \mid k \geq 0 \}$ 不是正规语言，试利用该结论以及正规语言封闭运算，证明如下语言 L 不是正规语言：

$$L = \{w \mid w \in \{a , b, c\}^*，\text{且 } w \text{ 中 } a \text{ 的个数比 } b \text{ 的个数多 } 2 \}$$

2. (5 分) 设 Σ 和 T 为字母表，以及映射 $h: \Sigma \rightarrow T^*$。对 $w=a_1a_2...a_n \in \Sigma^*$, 定义

$h(w) = h(a_1)\, h(a_2) \ldots h(a_n)$,　　称为串 w 的一个同态；

对语言 $L \subseteq \Sigma^*$, 定义 L 的同态 $h(L) = \{ h(w) \mid w \in L \}$。

我们有结论：若 S 为正规语言， 则 $h(S)$ 也是正规语言。

以下是该结论的一个证明过程：

证明　设 S 对应的正规表达式为 E ，使得 $L(E)=S$. 归纳于 E 的结构，可以证明: 存在正规表达式 $h(E)$ ，满足 $L(h(E)) = h(S)$.

基础：若 E 为 ε, ϕ, 取 $h(E) = E$, 显然 $L(h(E)) = h(L(E))$;

若 E 为a，取 $h(E) = h(a)$，有 $L(h(E)) = h(L(E)) = \{h(a)\}$;

归纳：若 $E=E_1E_2$，取 $h(E) = h(E_1)\,h(E_2)$，有

0） $L(h(E)) = L(h(E_1))\,L(h(E_2))$

1） $= h(L(E_1))\,h(L(E_2))$

2） $= h(\{w_1 \mid w_1 \in L(E_1)\})\,h(\{w_2 \mid w_2 \in L(E_2)\})$

3） $= \{h(w_1) \mid w_1 \in L(E_1)\}\,\{h(w_2) \mid w_2 \in L(E_2)\}$

4） $= \{h(w_1)h(w_2) \mid w_1 \in L(E_1) \wedge w_2 \in L(E_2)\}$

5） $= \{h(w_1w_2) \mid w_1w_2 \in L(E_1)L(E_2)\}$

6） $= h(L(E_1)L(E_2))$

7） $= h(L(E_1E_2))$

8） $= h(L(E))$

$E=E_1+E_2$ 和 $E=E_1{}^*$ 的情形类似，略。

试解释上述归纳步骤中，从 0）到 1）、从 2）到3）、从 3）到4）、从 4）到5）和从 6）到7）的理由。要求从以下可供选择的理由中找出最恰当的选择：

① 正规表达式语言的定义（即正规表达式连接运算的语义）

② 语言连接运算的定义

③ 语言同态的定义

④ 字符串同态映射的性质（即运算保持性）

⑤ 归纳假设

3. (5 分) 考虑由下列产生式定义的上下文无关文法 G：

$S \to a\,S\,b \mid S\,S \mid \varepsilon.$

试证明 $L(G) = \{\, w \mid w \in \{a, b\}^*$ ，w 中 a 和 b 的数目相同，且 w 的任意前缀中 a 的数目不少于 b 的数目$\}$