哈工大 2004 年 秋 季 学期

形式语言 试题

学号	0?3403328
姓名	[illegible]

题号	一	二	三	四	五	六	七	总分
分数								

（总分 90 分）

注意行为规范

遵守考场纪律

一、基本概念（每小题 4 分，共 16 分）

1. 简述有限状态自动机接受的语言与正则表达式之间的关系

1°. 对于有限状态的自动机所接受的语言可以用某一正则表达式来表达

2°. 对于某一正则表达式可以设计一个有限状态自动机来接受，且只接受这一正则表达式表示的语言。

2. 一个下推自动机 P 以终结状态方式接受的语言 L(P)的定义

设 $P=(Q,\Sigma,\Gamma,\delta,q_0,Z_0,F)$ 是一个 PDA，那么 P 以终结状态方式接受的语言 L(P) 是

$$\{w \mid (q_0,w,Z_0) \vdash_P^* (q,\varepsilon,\alpha)\}$$

其中 q 是 F 中的某个状态，α 是任何堆栈符号串，从以 w 为等待输入串的初始 ID 出发，P 消耗了输入的 w 并且进入了接受状态。在那一时刻栈中的内容无关。

3. DPDA 的形式定义

定义一个 PDA $P=(Q,\Sigma,\Gamma,\delta,q_0,Z_0,F)$ 为确定型的（称为确定型 PDA 或 DPDA）当且仅当下列条件被满足

1. 对于所有的 Q 中的状态 q，Σ 中的输入符号 a 或者 a=ε 以及 Γ 中堆栈符号 X，$\delta(q,a,X)$ 至多有一个成员。

2. 如果对于 Σ 中的某些输入符号 a 有 $\delta(q,a,X)$ 非空，那么 $\delta(q,\varepsilon,X)$ 一定

主管领导审核签字 张山石

4. M 为图灵机 TM，描述 M 所接受的语言

形式化的说，设 $M=(Q, \Sigma, \Gamma, \delta, q_0, B, F)$ 是一个图灵机，则 $L(M)$ 是对于属于 F 的某个状态 q 以及任意的带串 α 和 β，使得 $q_0 w \vdash^{*} \alpha p \beta$ 的 Σ^{*} 中串 w 的集合

Q：有穷控制的状态的集合　　F：终结状态，或接受状态

Σ：输入符号的有穷集合

Γ：带符号的完整集合

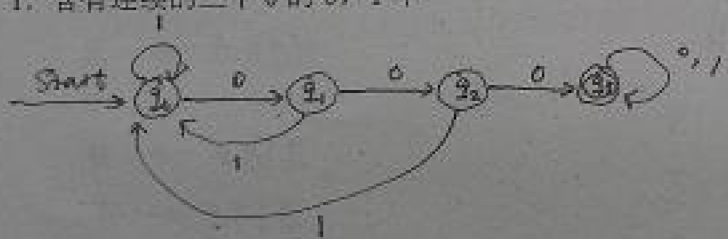
δ：转移函数

q_0：初始状态

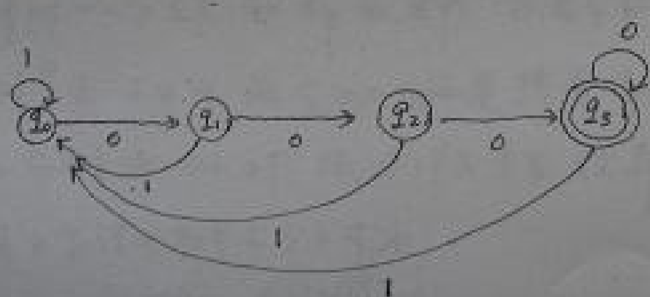
B：空格符号

二、构造 DFA（画出状态转移图），使之接受下列语言（20 分）

1. 含有连续的三个 0 的 0，1 串

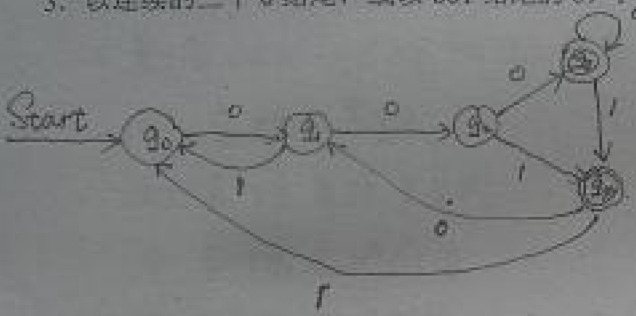


2. 以连续的三个 0 结尾的 0，1 串



x)一定为空。

3、以连续的三个 0 结尾，或以 001 结尾的 0，1 串



4、分别给出以上三种语言的正则表达式

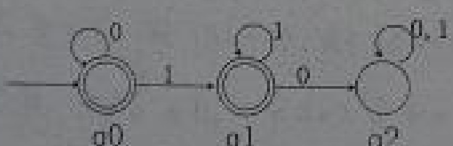
1. $(0+1)^* 000 (0+1)^*$

2. $(0+1)^* 000$

3. $(0+1)^* 00(0+1)$

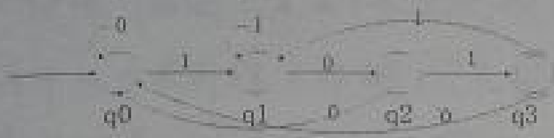
三、描述下列自动机所接受的语言（10 分）

1、



1. 不含 10 的 由 0，1 所构成的 0，1 串

2.



1. 以101结尾的0，1所构成的字符串.

四、设 $L=\{0^i1^j2^m3^n \mid i\neq m \text{ 或 } j=n\}$（14分）

1、写出 CFG G，使 L(G)=L

解：$S \rightarrow 0Q2F \mid G1P3 \mid \varepsilon$

$Q \rightarrow 0Q2 \mid 0A \mid C2$

$A \rightarrow 0A \mid B$

$C \rightarrow C2 \mid B$

$B \rightarrow 1B \mid \varepsilon$

$F \rightarrow 3F \mid \varepsilon$

$G \rightarrow 0G \mid \varepsilon$

$P \rightarrow 1P3 \mid R$

$R \rightarrow 2R \mid \varepsilon$

V: {S, Q, A, B, C, F, G, P, R}

T: {0, 1, 2, 3}

S: start symbol

2、判断上述 CFG 是不是二义的，并给出理由

这个文法一定是二义的，因由文法的产生式可以看出整体可以分作两部分：0Q2F 及 G1P3.
分别产生 $i\neq m$ 及 $j=n$ 两类符号串. 而 $j=n$ 且 $i\neq m$ 可以分别由 0Q2F 及 G1P3 产生，故存在这样的字符串分别有2个语法树与之对应.

故此文法歧义。

3. 将上述 CFG 转化为 CNF 形式

1. 遵循 CFG → CNF 的标准步骤

(a). 去除ε产生式

(b) 去除单位产生式

(c) 去除无用符号

(d) 转化为 CNF.

于是对应(a)有 $S \to 0Q2F|0Q2|G1P3|G1P|1P3|13$;

$Q \to 0Q2|0A|0|C2|2$,

$C \to C2|2|B$; $B \to 1B|1$; $F \to 3F|3$

$G \to 0G|0$ $P \to 1P3|13|2R|2$

$R \to 2R|2$

(b) 去单位对 {A,B} {C,B} {P,R}

(d) (0—H, 1—I, 2—K, B—L)

$S \to HC_1|HC_2|GC_3|GC_4|IG_5|IL$

$Q \to HC_2|HA|CK|0|2|$

$A \to HA|IB|0|1|$

$C \to CK|IB|1|2$

$B \to IB|1$

$F \to LF|3$

$G \to HG|0$

$P \to IG_5|KR|IL|2$

$R \to KR|2$

$C_1 \to QC_6$

$C_2 \to QK$

$C_3 \to IC_5$

$C_4 \to IP$

$C_5 \to PL$

$C_6 \to KF$

$H \to 0$

$I \to 1$

$K \to 2$

$L \to 3$

五、试用相应的泵引理证明（15 分）

1. $L=\{0^n \mid n$ 为完全平方数$\}$不是正则语言

: 假设 L 是正则语言 取串 $w=0^{n^2} \in L$

$\geq n$ (n 为泵引理常数)

令 w 被打断为三段 $w=xyz$ 其中

$|xy| \leq n$ 且 $y \neq \varepsilon$ 设 $|y|=m$ 则 $1 \leq m \leq n$

由泵引理有 $xy^kz \in L$ ($k \geq 0$) 令 $k=2$

$|xy^2z| = n^2+|y| = n^2+m$ 显然

$n^2+1 \leq n^2+m \leq n^2+n < n^2+2n+1=(n+1)^2$

在 $(n, n+1)$ 区间内找到一个整数,

其平方满足题中要求 不可能.

L 不是正则语言

2. $L=\{a^i b^j c^k \mid i<j<k\}$ 不是上下文无关的

证明：假设L是上下文无关的，取串 $z=a^n b^{n+1} c^{n+2} \in L$，$|z| \ge n$

其中n为泵引理之常数，于是将其分成五段 $z=uvwxy$ 其中 $|vwx| \le n$

且 $vx \ne \varepsilon$，∵ $|vwx| \le n$ 故 vwx 中不可能同时包含 a 和 c，vwx

的字符串显然只分为以以下几种情况：

1. vwx 只包含 a，由 $vx \ne \varepsilon$，故a至少一个，则串 uv^2wx^2y 中a的个数 $|a| \ge n+1$
   $|b|=n+1$，∴与题中條件矛盾
2. vwx 只包含至少一个 b 那么 uv^2wx^2y 中 $|b| \ge n+2$；$|c|=n+2$ ∴矛盾
3. vwx 只包含有至少一个c，那么 $|c| \le n+1$，而 $|b|=n+1$（uwy）∴矛盾
4. vwx 只包含有至少一个a和至少一个b，uv^2wx^2y 中 $|b| \ge n+2 = |c|$ 矛盾
5. vwx 只包含有至少一个b和至少一个c，则串 uwy 中 $|b| \le n \le |a|$ ∴矛盾

由上述分析知，L不可能是上下文无关的。

六、写出与下列PDA等价的CFG，要求非终结符采用[pxq]形式（10分）

$\delta(q,0,1)=\{(q,\varepsilon)\}$; $\delta(q,1,Z)=\{(q,1Z)\}$; $\delta(q,0,Z)=\{(q,0Z)\}$;

$\delta(q,\varepsilon,Z)=\{(q,\varepsilon)\}$; $\delta(q,1,0)=\{(q,\varepsilon)\}$;

$\delta(q,1,1)=\{(q,11)\}$; $\delta(q,0,0)=\{(q,00)\}$;

解：PDA $P=\{\{q\},\{0,1\},\{0,1,Z\},q,Z,F\}$

于是 CFG G 中 V 有4个 S [q0q], [q1q], [qZq]

产生式为

$S \to [qZq]$

$[q1q] \to 0$

$[qZq] \to 1[q1q][qZq]$

$[qZq] \to 0[q0q][qZq]$

$[qZq] \to \varepsilon$

$[q0q] \to 1 \mid 0[q0q][q0q]$

$[q1q] \to 1[q1q][q1q]$

∴ $G=(\{S, [q0q][q1q],[qZq]\}, \{0,$

七、证明一个正则语言与一个 CFL 的交是 CFL（5 分）

证明：构造 PDA 来接收一正则语言与 CFL 的交

设 CFL 对应于 PDA $P = \{Q_p, \Sigma_p, \Gamma_p, \delta_p, q_p, Z_0, F_p\}$

正则语言对应于 $(Q_D, \Sigma_D, \delta_D, q_D, F_D)$

则构造一个 PDA 如下：

$(Q_p \times Q_D, \Sigma_p \cup \Sigma_D, \Gamma_p, \delta, (q_p, q_D), Z_0, F_D \times F_p)$

$\delta((q_p, q_D), a, \alpha) = ((q_{p_1}, q_{D_1}), \beta)$

其中 $\delta(q_p, a, \alpha) = (q_p, \beta)$

$\delta(q_D, a) = q_{D_1}$

由此所得 PDA 设计完毕，显然只接受且接受这一正则语言与一个 CFL 的交。