形式语言与自动机2003

图都没画。

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》 B卷 2003年 7 月 2 日

学号： 姓名：

1. (34 pts) 回答 true 或者 false ( Answer true or false ): (2 pts. each)

a) \_\_\_\_\_ 如果一个语言可由某个右线性文法生成，那么它是正规的 （If a language can be generated by a right-linear grammar, it is regular） .

b) \_\_\_\_\_ 对于每一个可以被确定 PDA 接受语言，都存在一个接受它的非确定的 PDA （For every language accepted by a deterministic PDA, there is a nondeterministic PDA that accepts the same language）.

c) \_\_\_\_\_\_ 每个二义文法可以转化为等价的无二义文法（Every ambiguous grammar can

be converted to an equivalent unambiguous grammar）.

d) \_\_\_\_\_\_ 所有C++程序的集合是一个形式语言（The set of all possible C++ programs is a formal language）.

e) \_\_\_\_\_\_ 由以下产生式定义的文法符合 Chomsky 范式的形式（The following gram

mar is in Chomsky Normal Form）. S -> AB

A ->AA | a

B ->BB | b

f) \_\_\_\_\_\_\_ 针对正规语言的 pumping 引理不可用于证明一个语言是正规的（The pumping lemma for regular languages can't be used to prove that a language is regular）.

g) \_\_\_\_\_\_\_ 语言 L= { wwR|w属于 {a, b}\* } 不是正规的，其中 wR 代表 w 的反向字符串（下同）

（The language L= { wwR|w 属于 {a, b}\* } is not regular，where wR is the revers

al of w ( the same in the rest part below )）.

h) \_\_\_\_\_\_\_ 对一个 DFA 应用填表算法，合并所有不可区分的状态，则可获得一个等价于该 DFA 的拥有"最少"状态数目的DFA（After applying the table-filling algorithm to a DFA and combining all the distinguished states, one can get an equivalent DFA with the least state number）.

注：此题选择False

i) \_\_\_\_\_\_\_ 对每个可以被 DPDA 接受的语言，都存在一个无二义的上下文无关文法接受它（Every language accepted by a DPDA has an unambiguous context-free grammar）.

j) \_\_\_\_\_\_\_ 递归语言是递归可枚举语言（A recursive language is a recursively enumerable language）.

k) \_\_\_\_\_\_\_ 一个语言可被带三个栈的 PDA 接受，但它不可以由任何带两个栈的 PDA 接受（Every language accepted by a PDA with three stacks can't be accepted by any PDA with two stacks）.

l) \_\_\_\_\_\_\_ 一个语言可被带两个计数器的计数器机接受，则它也可由带一个计数器的计数器机接受（Every language accepted by a two-counter machine is also accepted by a one-counter machine）.

m) \_\_\_\_\_\_\_ 对任何语言 L1 和 L2，如果 L1是正规的，那么 L1 U L2 是正规的（For any languages L1, L2, if L1 is regular, then L1 U L2 is regular）.

n) \_\_\_\_\_\_\_ 对任何语言 L1 和 L2，如果 L1 和 L2 不是正规的，那么 L1 交 L2 不是正规的 （For any languages L1, L2, if L1 and L2 are not regular, then L1 交 L2 is not regular）.

o) \_\_\_\_\_\_\_ 对任何正规语言 L1 和 L2，L1 U L2 L1 =L1 (L1 U L2)（For any regular languages L1 and L2, L1 U L2 L1 = L1 (L1 U L2)）.

p) \_\_\_\_\_\_\_ 对任何上下文无关语言 L1 和 L2，L1 - L2 也是上下文无关的（For any context-free languages L1 and L2, L1 - L2 is context-free）.

q) \_\_\_\_\_\_\_ 终态方式接受的PDA所接受的语言，一定可以被空栈接受方式的PDA所接受；类似地，终态方式接受的DPDA所接受的语言，一定可以被空栈接受方式的DPDA所接受.（A language accepted by final state by a PDA is accepted by empty stack by a PDA. Si

milarly, a language accepted by final state by a DPDA is accepted by empty s

tack by a DPDA）.

2. (12 pts) 考虑正规表达式r = b(a + b)a\*（Consider the regular expression r = b(a + b)a\*）.

a) 构造一个接受L(r) 的e-NFA（Construct a e-NFA that accepts L(r)）. (3 pts)

b) 给出语言L(r) 的反向的一个正规表达式（Give a regular expression for the reversal of L(r)）. (3 pts)

c) 构造可以生成语言 L(r) 的一个上下文无关文法（Construct a context-free grammar generating the language L(r)）. (6 pts)

3. (6 pts) 下图是一个扩展的两个状态的有限自动机，其中 R，S，T，和 U 是正规表达式。可以用不同的形式给出等价于该扩展自动机的正规表达式，试给出两种这样的正规表达式。 （Give two regular expressions which describe the following generic two-state finite automata in two different ways, where R, S, T, and U are regular expressions.）.

4. (6 pts) 正规表达式 ab\* 表示 {a, b} 上的一个语言。试构造接受该语言的补语言的

一个 DFA。（The regular expression ab\* denotes a language on {a, b}. Constru

ct a DFA that accepts the complement of this language.）

5. (8 pts) 试给出下列每个正规语言的一个正规表达式（Give a regular expression for each of the following regular languages）：

a) { xwxR | x, w属于 (a + b)+ }, where (a + b)+ = (a + b) (a + b)\*

b) { w |w 属于 {a , b}\* 且 存在x, y( x, y 属于{a , b}且 w = xy 且 | x | =3 且 x = xR ) }

最后这个R是x的上标

6. (7 pts) 下面是图灵机 M 的转移图。对于输入字符串01122，给出 M 从初始 ID 开始可以到达的所有可能的 ID （Following is a transition diagram for Turing machine M.

Show all of the ID's reachable from the initial ID if the input is 01122）.

7. (8 pts) 考虑由下列产生式定义的上下文无关文法 G（Consider the context-free

grammar G defined by productions）：S->0S1 | e. （e是那个表示空的希腊字符）

证明L(G)={ 0n1n | n大于等于0} （ 0n,1n分别表示0，1的n次方）

（Prove that L(G)={ 0n1n | n大于等于0}）.

8. (8 pts) 下列文法可产生正常嵌套的括号串，所有正常嵌套的括号串构成一个语言。试

构造一个接受该语言的空栈接受方式的下推自动机（6 points），然后再将其转换为等

价的终态接受方式的下推自动机（2 points）（The following grammar generates all strings of properly nested parentheses. Construct a pushdown automaton that accepts this language by empty stack and then convert it to an equivalent

PDA which accepts the same language by final sate）.

S->(S）| SS | e. (e是那个表示空的希腊字符）

9. (5 pts) 下列文法可产生所有正常嵌套的括号串。试消去该文法中的 e-产生式，使得

到的新文法可以产生除空串之外的所有正常嵌套的括号串（The following grammar ge

nerates all strings of properly nested parentheses. Eliminate the e-producti

ons to form a grammar generating the nested parentheses without empty string

）.

S->（S）S | e.

10. (6 pts) 应用针对上下文无关语言的 pumping 引理，

说明语言L={akbk+1ak| k大于等于0} （ak为a的k次方，bk+1为b的k+1次方）

不是上下文无关的 （Use the pumping lemma for context-free languages to show the language L = {akbk+1ak| k>=0} is not context free）.

--

形式语言2002版(不完全)

1. 多项选择题

(1) a\*

(2) {a^nb^n}

(3) {a^nb^nc^n}

(4) {a^nb^mc^n}

(5) 对角化语言

(6) 通用语言

A. 有穷自动机接受

B. 下推自动机接受，但是不能被有穷自动机接受

C. 有穷自动机接受, 不能被空栈接受方式的DPDA接受

D. 可被一定会停机的图灵机接受，不能被有穷自动机接受

E. 可被图灵机接受，不能被一定会停机的图灵机接受

F. 不能被图灵机接受

参考答案：(1) AC (2) B (3) D (4) B (5) F (6) E

2. 图灵机当前ID是: (q0)0110

转移方式是:

delta(q0,0)=(q1,1,R)

delta(q1,1)=(q2,0,L)

delta(q2,1)=(q0,1,R)

delta(q2,B)=(q3,B,R)

q3是终止状态

参考答案：到111(q2)0停机，没有接受

3. 设有四个语言（或问题）A,B,C,D。这些语言可以是，也可以不是递归

可枚举语言，但已知如下条件：

i A 可以规约到 （不一定是多项式时间规约，下同）

ii B 可以规约到 C

iii D 可以规约到 C

对于以下6个命题，分析指出它们是“真”，或是“假”，或是“不能确定”

(a)A是递归可枚举的但不是递归的，而C是递归的。

(b)B不是递归的，而D是递归可枚举的。

(c)A的补不是递归可枚举的，而B的补是递归可枚举的。

(d)B的补不是递归的，而C的补是递归的。

(e)如果A是递归的，那么B的补是递归的。

(f)如果C是递归的，那么D的补是递归的。

4. CYK算法

5. 判断题，判断有穷自动机和下推自动机是否可判定

(a) 判定是否为空

(b) 判定s是否属于

(c) 判定两个是否相同

最后判定递归语言和递归可枚举语言是否满足(b)

参考答案：(a)可判定 (b)可判定 (c) 有穷自动机可判定，另一个不可判定

递归语言可判定，递归可枚举语言不可判定

6. 证明S->ASB|AB,A->a,B->b等价于{a^nb^n,n>=1}

7. 通过封闭运算证明{a^nb^(n+2)}是CFL,证明{a^nb^ic^(n-i)d^je^(n-j)}不是CFL

8. 判断两个语言是否相等，或者某个包含于某个

(1) (a) S->1S2|2S1|epsilon

(b) (1+2)\*

(2) (a) 一个CFL

(b) {0^n1^n,n>=0}

(3) (a) (0+1)\*11(0+1)\*

(b) (0\*1\*11)\*0110\*

(4) (a) 一个CFL

(b) {0^i1^j,i>=0,j>=0}

(5) (a) 一个CFL

(b) S->.......

参考答案：(1) A<B (2) A<B (3) A>B (4) A=B (5) A=B (顺序可能不一致)

9. 选择题

选项:

REG, CFG, NONE

(1) 0,1字符串，任何前缀的0,1字符个数的差绝对值小于2

(2) 0,1字符串，0的个数比1的个数多2

(3) 0,1,2字符串，0的个数比1,2个数总数要多

(4) 0,1,2字符串，0的个数多于1的个数，且多?的个数

(5) 至少包含100个0,100个1,100个2的0,1,2字符串

参考答案：(1) REG (2) CFG (3) CFG (4) NONE (5) REG

--

自动机2004(A卷)

一、选择题（次序有问题） (12pts)

1.00结尾的串

2.L={0n1n | n>=0}

3.L={0i1j | i<=j}

4.L={0n1n2m | n>=m}

5.wwR(wR表示w的取反)

6.a\*b

A 是正规语言，也是某个空栈DPDA

B 是正规语言，不是某个空栈DPDA

C 是终态接受的DPDA, 不是空栈接受的DPDA

D 是终态接受的DPDA, 也是空栈接受的DPDA

E 是PDA，但不是DPDA

F 是图灵机，不是下推自动机

二 、判断题 (16pts)

1.L1,L2是正规语言，则(L1交L2)^\*(L1并L2)^+是正规语言．（大概是这样的形式）

2.L1是正规语言，L1包含于L2, 则L2是正规语言

3.L1是正规语言，L2包含于L1, 则L2是正规语言

4.存在一个算法可以判定PDA是不是DPDA

5.如果G是Chomsky范式，那么Epsilon不属于G

6.L1是CFG, L2是正规语言，则L1-L2是CFG

7.NP-complete和NP-HARD都是NP问题

8.递归可枚举语言是递归语言

三、简答题 (32pts)

1. 根据DFA，写一个RE，RE是0\*(1+e)，图很简单

2.

3.S->SaS|bSc|\Epsilon

改写成无歧异文法

4.给一个DFA，要画出它反同态的DFA

h(0)=ab h(1)=ba

原来的DFA: 转移函数写成f

f(q0,a)=q1 f(q0,b)=q2

f(q1,a)=q1 f(q1,b)=q0

f(q2,a)=q1 f(q2,b)=q2

5.如何判断DFA A的语言L(A)是不是等于sima\*，用自然语言描述

6.消去单一产生式

S->A|B

A->B|SS

B->(S)|()

7.

8.简述用有2个计数器的计数器机接受语言0n1n2n

四、 (24pts)

1.L={(a+b)\*，且没有连续的aa}，写出DFA

2.L={(a+b)\*，且没有连续的aa,bb} 写出正规表达式

3.给一个CFG，接受{a^nb^mc^id^j|n+m=i+j}

4.图灵机，输入N的二进制表示，如果N为奇数，把它变成2N+1，如果N为偶数，把它变成N+1

五、证明题 (14pts)

1.(6pts)用pumping引理证明L={a^m b^n c^p | m<n<p, 且m,n,p是自然数}不是CFG

2.(4pts)

G:

S->aSc|A

A->bAc|B

B->Bc| \Epsilon

证明: L(G) = {a^i b^j c^k | i+j<=k, i,j,k是自然数}

3.(4pts) CFG G: S->SaS|bSc| \Epsilon

证明：它不是正规语言。

2002年的题，王生原老师放在网络学堂上的，更完全，表述更标准一些，也留给以后的ddmm吧

《形式语言与自动机》期末测试题（规定时间：2小时）

学号 姓名

1. (12 pts) 选择填空（多选）：

(a) 由正规表达式 a\* 定义的语言 ；

(b) 语言 {a^m\*b^m|m ≥0} ；

(c) 语言 {a^m\*b^m\*c^m|m ≥0} ；

(d) 语言 {a^m\*b^n\*c^m|m,n ≥0} ；

(e) 对角化（diagonalization）语言 Ld ；

(f) 通用语言（universal）语言 Lu ；

供选择的答案：

A. 是某个有限状态自动机的语言；

B. 是某个下推自动机的语言, 但不是任何有限状态自动机的语言；

C. 是某个有限状态自动机的语言,但不是任何空栈接受方式的确定下推自动机的语言；

D. 是某个可停机的图灵机的语言, 但不是任何下推自动机的语言；

E. 是某个图灵机的语言, 但不是任何可停机的图灵机的语言； F. 不是任何图灵机的语言.

2. (7 pts) 图灵机（{ q0, q1, q2, q3 }, { 0,1 }, { 0,1,B }, δ, q0, B, {q3 }）有如下转移规则：

δ（q0, 0）=（q1,1, R）

δ（q1, 1）=（q2, 0, L）

δ（q2, 1）=（q0, 1, R）

δ（q1, B）=（q3, B, R）

给出从初始ID q00110开始, 该图灵机可以到达的完整的ID序列, 直到它停机.

3. (12 pts) 设有四个语言（或问题）A，B，C 和 D. 这些语言可以是，也可以不是递归可枚举语言，但已知如下条件：

i. A 可以归约到 B（不一定是多项式时间归约，下同）；

ii. B 可以归约到 C；

iii. D 可以归约到 C；对于以下 6 个命题，分别指出它们是“真”，或是“假”，或是“不能确定”：

(a) A 是递归可枚举的但不是递归的，而 C 是递归的.

(b) B 不是递归的，而 D 是递归可枚举的.

(c) A 的补不是递归可枚举的，而 B 的补是递归可枚举的.

(d) B 的补不是递归的，而 C 的补是递归的.

(e) 如果 A 是递归的，那么 B 的补是递归的.

(f) 如果 C 是递归的，那么 D 的补是递归的.

4. (11 pts) 文法 G 的产生式集合为： S→AS|b, A→BB|a|b, B→BA|a. 下面的左图表示对于文法 G 和字符串 ababb 应用 CYK 算法时所构造的表，各集合与表元素符号之间的对应关系可参照右边的图. 试问

(a) 计算 X15 时与哪几对表元素相关？ (5 pts)

(b) X15 = ？ (5 pts)

(c) 是否有 ababb ∈ L(G)? (1 pt)

| X15 | X15

| {S,A,B} {S} | X14 X25

| {A,B} {S} {S,B} | X13 X24 X35

| {S,B} φ {S,B} {S} | X12 X23 X34 X45

| {A,B} {S,A} {A,B} {S,A} {S,A} | X11 X22 X33 X44 X55

————————————————— —————————————

a b a b b a1 a2 a3 a4 a5

5. (8 pts) 对于正规语言和上下文无关语言两类语言，分别指出下列四个问题是否可判定的：

(a) 给定该语言类中的任一语言 L，判定 L 是否为空？

(b) 给定该语言类中的任一语言 L 和任何字符串 w，判定 w 是否属于 L？.

(c) 给定该语言类中的语言 L 和 M，判定 L 是否与 M 相等？.

对于递归语言和递归可枚举语言两类语言，分别指出上述问题（b）是否可判定的：

6. (15 pts) 对下面描述每一对语言，指出以下 4 种选择中的哪一个正确表达了它们之间的关系：

（1）包含于（2）：语言（1）是语言（2）的真子集；

（1）包含（2）：语言（2）是语言（1）的真子集；

（1）=（2）：语言（1）和语言（2）相等；

（1）？（2）：语言（1）和语言（2）之间没有包含关系，即（1）中的某些串不在（2）中，（2）中的某些串不在（1）中.

(a) （1）产生式集为 S→0S1 | 1S0 | ε 的 CFG 的语言.

（2）正规表达式 (0+1)\* 的语言.

(b) （1）以终态方式接受的 PDA （{ p，q }，{ 0,1 }，{ X,Z }，δ，q，Z，{ p }）的语言，其中 δ 转移包括：

δ(q,0,Z)=(q,XZ), δ(q,0,X)=(q,XX), δ(q,1,X)=(p, ε) δ(p,1,X)=(p, ε).

（2）产生式集为 S→0S1 | 0S | ε 的 CFG 的语言.

(c) （1）产生式集为 S→AS|SB|ε，A→0 和 B→1 的 CFG 的语言.

（2）正规表达式 0\*1\* 的语言

.

(d) （1）正规表达式 (0+1)\*11(0+1)\* 的语言.

（2）正规表达式 (0\*1\*11)\*0\*110\*1\* 的语言.

(e) （1）NFA （{ p，q }，{ 0,1 }，δ，q，{ q }）的语言，其中转移函数 δ为：

δ(q,0)={p }, δ(q,1)=ε, δ(p,0)={p,q}, δ(p,1)={p}.

（2）产生式集为 S→AS|ε，A→0B 和 B→0B|1B|0 的 CFG 的语言.

7. (10 pts) 利用上下文无关语言的闭包运算可以证明某些语言是或不是上下文无关语言.

试证明 {a^n\*b^(n+2)|n ≥0} 为上下文无关语言，而 {a^n\*b^(n-i)\*c^i\*d^(n-j)\*e^j |i,j ≤n} 不是上下文无关语言（提示，已知 {0^n1^n|n ≥0} 为上下文无关语言，而 {0^n1^n2^n |n ≥0} 不是上下文无关语言）.

8. (10 pts) 设上下文无关文法G =（{S},{a,b},P,S），其中P包含如下产生式：S－>aSb|ab .

试证明 L(G)={a^nb^n|n ≥1}.

9. (15 pts) 对下面描述每一个语言，指出以下 4 种选择中的哪一个正确描述了该语言：

REG：该语言是正规的；

CFG：该语言是上下文无关的但不是正规的；

NON：该语言不是上下文无关的；

(a) 满足如下条件的 0，1 字符串集合：任何一个前缀所包含的"0"的数目与所包含的"1”的数目相差不超过 2 （可能 "0" 比"1" 多，也可能 "1"比"0" 多）.

(b) 满足如下条件的 0，1 字符串集合：所包含的"1" 的数目恰好是所包含的"0"的数目多 2.

(c) 满足如下条件的 0，1 ，2 字符串集合：所包含的"0" 的数目是最多的.

(d) 满足如下条件的 0，1 ，2 字符串集合：所包含的"0" 的数目大于 "1" 和 "2"的数目之和.

(e) 满足如下条件的 0，1 ，2 字符串集合：所包含的"0"，"1" 和 "2"的数目至少都为 100.