HKUST – Department of Computer Science and Engineering

COMP 3721: Theory of Computation – Spring 2015

Serial #

Midterm Examination

Name: ____________________ Student ID: ____________________

Instructions

- This is an open-book and open-notes examination. Computers and smart phones are not allowed.
- They are 8 questions and 13 pages. You can write on the back of the pages if necessary. The last three pages can be torn off and used as scrap papers.Let us know if you need more papers.
- Write your name and student ID on this page. For each subsequent page, write your student ID at the top of the page in the space provided.
- When asked to construct an automaton, you can use either a state diagram (recommended) or the formal mathematical definition.

Questions	1	2	3	4	5	6	7	8	Total
Score									

Student ID: ______________

Problem 1: [10 pts] Let N be the set of natural numbers. Prove that the following set (Cartesian product of countably many copies of N) is uncountable:

$$S = N \times N \times N \times \dots$$

Proof:

Assume the S is countable. Enumerate **all** its elements as V_1, V_2, ..., where each element V_i is an infinite vector of natural numbers:

$$V_i = (n_{i1}, n_{i2}, \dots)$$

Define a vector $V = (v_1, v_2, \dots)$ as follows:

$$v_i = \begin{cases} 0 & \text{if } n_{ii} \neq 0 \\ 1 & \text{if } n_{ii} = 0 \end{cases}$$

Then V differs from each V_i at the i-th element. Hence it is not in the enumeration. This is a contradiction. Therefore, the set S is uncountable.

Student ID: ____________

Problem 2: [10 pts]

(a) Give a regular expression for the language

$$L = \{0^n 1^m : n \geq 4 \text{ and } m \leq 3\}.$$

(b) Give a regular expression for the complement $\bar{L}$ of the language L.

Solutions:

(a) $00000^*(e \cup 1 \cup 11 \cup 111)$.

(b) $(e \cup 0 \cup 00 \cup 000)1^* \cup 0^*11111^* \cup (0 \cup 1)^*10(0 \cup 1)^*$

Grading: a: 5; b: 5

Student ID: ____________

Problem 3: [16 pts]

Consider the following language:

$L = \{w \in \{0,1\}^* : w$ is the binary representation of a natural number divisible by $4\}$.

For simplicity, assume that a binary number can begin with '0'. For example, '010' and '10' are both legal representations of the natural number 2.

(a) Write a regular expression for the language L.

(b) Construct an NFA that accepts L by following the steps described in Lecture 6.

(c) Convert the NFA of part (b) into a DFA by following the steps described in Lecture 6.

[MORE WORKSPACE ON NEXT PAGE]

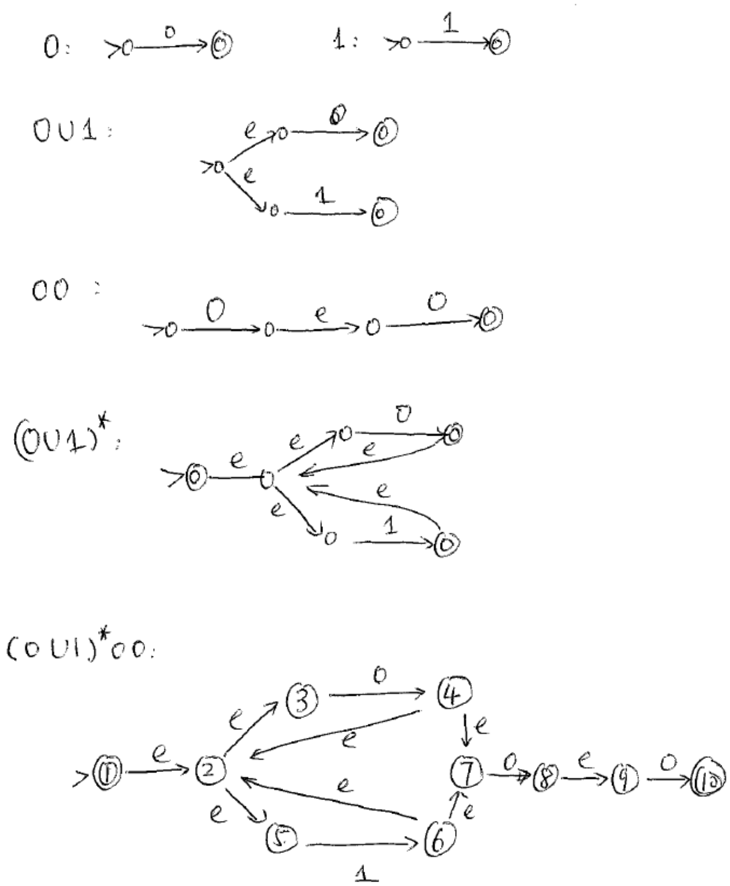
Solutions :

(a) $(0 \cup 1)^*00$.

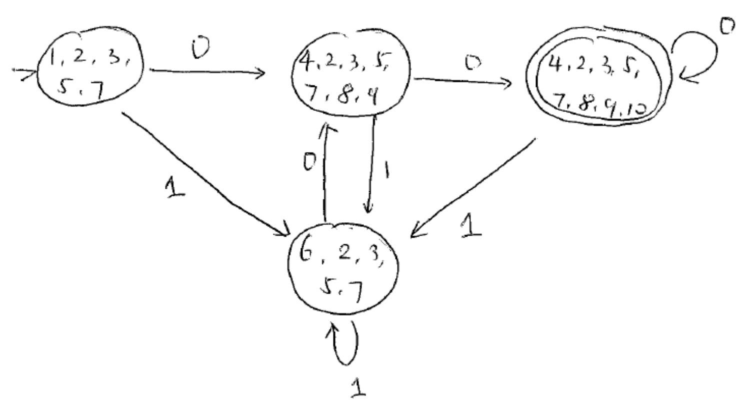
Grading: a:4; b:6; c:6

[WORKSPACE FOR PROBLEM 3]

(b)



(c)



Problem 4: [15 pts] Is each of the following languages regular? If it is, write a regular expression for it. If it is not, give a proof.

(a) $L = \{wa^n b^n : w \in \{a, b\}^*, n > 0\}$.

(b) $L = \{wa^n b^n : w \in \{a, b\}^*, n \geq 0\}$.

Proof :

(a) The language is not regular. We prove the statement by contradiction. Assume L is regular. Then the following intersection should also regular:

$$L_1 = L \cap L(a^* b^*) = \{a^m b^n : w \in \{a, b\}^*, m \geq n > 0\}.$$

In the following, we show that L_1 is actually not regular. Hence, L cannot be regular.

We show that L_1 is not regular again by contradiction. Assume it is regular. Let N be the integer from the Pumping Theorem for regular languages. Consider $w = a^N b^N \in L_1$. According to the Pumping Theorem, w can be written as:

$$w = xyz, \text{ s.t. } y \neq e, |xy| \leq N, xy^i z \in L_1 \ \forall i = 0, 1, \ldots.$$

Because $y \neq e, |xy| \leq N$, $y = a^k$ for some $k > 0$. Therefore, $xy^0 z$ would have fewer a's than b's, and hence cannot be in L_1. This is a contradiction. So, L_1 cannot be regular.

The proof is completed.

(b) The language is regular. Actually, $L = \{w : w \in \{a, b\}^*\} = L((a \cup b)^*)$.

Grading: a: 10; b: 5.

Student ID: ______________

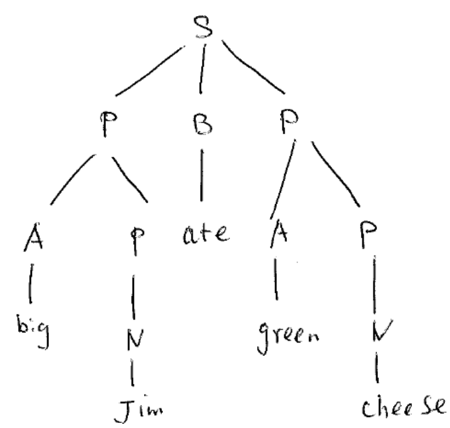
Problem 5: [6 pts]

Let G be the context-free grammar (V, Σ, R, S) where

$$\begin{aligned}
\Sigma &= \{\text{Jim, big, green, cheese, ate}\}, \\
V &= \{S, A, N, B, P\} \cup \Sigma, \\
R &= \{S \to PBP, P \to N, P \to AP, \\
&\quad A \to big, A \to green, N \to cheese, \\
&\quad N \to Jim, B \to ate.\}
\end{aligned}$$

Draw a parse tree for the string "big Jim ate green cheese".

Solution:



Problem 6: [15 pts]

Let $V = \{a, b, S, A, B, C\}$ and $\Sigma = \{a, b\}$. What language does each of the following context-free grammars generate?

(a) $G_1 = (V, \Sigma, R_1, S)$, where

$$R_1 = \{S \to A, A \to CAC, A \to a, C \to a, C \to b\}.$$

(b) $G_2 = (V, \Sigma, R_2, S)$, where

$$R_2 = \{S \to B, B \to CBC, B \to b, C \to a, C \to b\}.$$

(c) $G_3 = (V, \Sigma, R_3, S)$, where

$$R_3 = R_1 \cup R_2$$

(d) $G_4 = (V, \Sigma, R_4, S)$, where

$$R_4 = R_3 \cup \{S \to AB\}$$

Solutions :

(a) $L(G_1) = \{xay : x, y \in \{a, b\}^*, |x| = |y|\}$
(b) $L(G_2) = \{xby : x, y \in \{a, b\}^*, |x| = |y|\}$
(c) $L(G_3) = L(G_1) \cup L(G_2) = \{w \in \{a, b\}^* : \ w \text{ is of odd length}\}$
(d) $L(G_4) = L(G_1)L(G_2) \cup L(G_3)$.

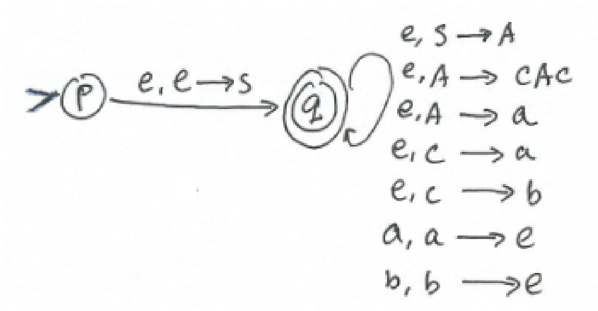
Grading: a+b: 5; c: 5; d: 5.

Problem 7: [10 pts]

(a) Construct a pushdown automaton that accepts the language $L(G_1)$ generated by the context-free grammar G_1 of Problem 6.

(b) Show the steps that the machine takes to accept the string *ababb*.

Solution: (a) Here is the PA:



(b) The steps that the machine takes to accept the string *ababb* are as follows:

$$
\begin{aligned}
(p, ababb, e) &\vdash (q, ababb, S) \vdash (q, ababb, CAC) \vdash (q, ababb, aAC) \\
&\vdash (q, babb, AC) \vdash (q, babb, CACC) \vdash (q, babb, bACC) \\
&\vdash (q, abb, ACC) \vdash (q, abb, aCC) \vdash (q, bb, CC) \\
&\vdash (q, bb, bC) \vdash (q, b, C) \vdash (q, b, b) \\
&\vdash (q, e, e).
\end{aligned}
$$

grading: a:5; b:5

Student ID: ____________

Problem 8: [18 pts] Is each of the following languages context-free? If it is, write a grammar for it. If it is not, prove your claim.

(a) $L = \{a^m b^n c^p : m, n, p \geq 0 \text{ and } m+n = p\}$.

(b) $L = \{a^m b^n c^p : m, n, p \geq 0 \text{ and } m \geq n \text{ and } m+n = p\}$.

Solution: (a) The language L in this part is context-free. It is generated by the following grammar rules:

$$S \to e, S \to A, A \to aAc, A \to B, B \to bBc, B \to e.$$

(b) The language L in this part is not context-free.

We will prove the fact by contradiction. Assume L is context-free. Let N be integer from the Pumping Theorem for context-free languages. Consider $w = a^N b^N c^{2N}$.

According to the Theorem, we can write w as follows:

$$w = uvxyz, \text{ s.t. } vy \neq e, |vxy| \leq N, uv^i xy^i z \in L \ \forall i = 0, 1, \ldots.$$

There are several cases:

* If the substring vxy are entirely from the first half of $a^N b^N c^{2N}$: In this case, uv^0xy^0z has $2N$ c's and the total number of a's and b's is smaller than $2N$. Hence, uv^0xy^0z does not belong to in L. A contradiction.
* If the substring vxy are entirely from the second half of $a^N b^N c^{2N}$: In this case, uv^0xy^0z has fewer than $2N$ c's and the total number of a's and b's is $2N$. Hence, uv^0xy^0z does not belong to L. A contradiction.
* If the substring vxy are from the end of the b-block and the beginning of the c-block: In this case, uv^ixy^iz has more b's than a's when i is larger enough, and hence does not belong to L. A contradiction.

Therefore, L is not context-free.

Grading: a: 7; b: 11