UNIVERSITY OF BRISTOL

Practice Examination Period

FACULTY OF ENGINEERING

Second Year Examination for the Degrees of Bachelor of Science Master of Engineering

COMS20007J
Programming Languages and Computation

TIME ALLOWED:
3 Hours

This paper contains *three* questions, worth *45*, *35* and *20* marks respectively. Answer *all* questions. The maximum for this paper is *100 marks*. Credit will be given for partial answers.

PLEASE WRITE YOUR 7 DIGIT STUDENT NUMBER (NOT CANDIDATE NUMBER) ON YOUR ANSWERS. YOUR STUDENT NUMBER CAN BE FOUND ON YOUR UCARD.

Other Instructions:

You may use the unit reference material without citation. You should not require any other material, but any other material that you do use must be cited clearly and correctly. You must not collaborate.

YOU MAY START IMMEDIATELY

Q1. This question is about regular languages.

(a) Consider each of the following regular expressions over the alphabet $\Sigma = \{a, b, c\}$:

1. $((b|c)^*a(b|c)^*a(b|c)^*)^*$
2. $\Sigma^*abb\Sigma^*$
3. $\Sigma^*(abb \mid baa)\Sigma^*$
4. $b\Sigma^*b$
5. $(\Sigma\Sigma)^*$
6. $(b|c)^*$
7. $abb\Sigma^*$

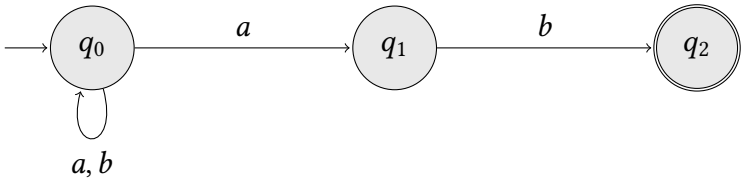
Match each of the following descriptions of languages to the regular expression above that denotes it:

i. The language of all words that start and end with b
ii. The language of all words that contain abb as a substring
iii. The language of all words that start with abb.
iv. The language of all words that do not contain a.
v. The language of all even length words.
vi. The language of all words containing an even number of a.
vii. The language of all words that either contain abb or bba

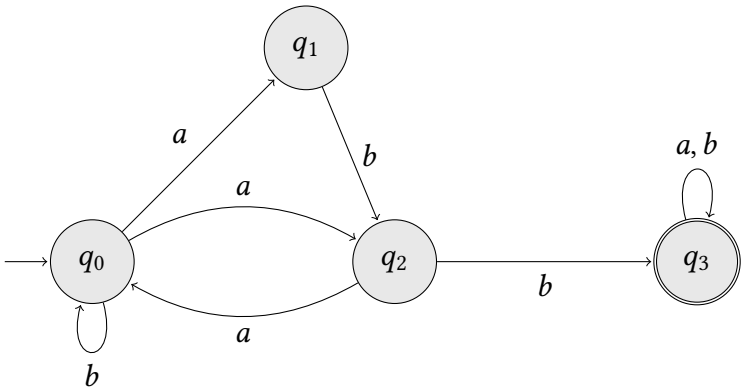
[7 marks]

(b) For each of the following automata, give a word that is accepted by the automaton *and* a deterministic automaton that recognises the same language.

i.



ii.



[10 marks]

(c) As in Problem Sheet 2, Question 7, we shall encode pairs of natural numbers by sequences of vectors of bits, with least significant bit first (left-most).

Let Σ be the following set of binary vectors:

$$\binom{0}{0}, \binom{0}{1}, \binom{1}{0}, \binom{1}{1}$$

We use each word w over alphabet Σ to encode a pair of natural numbers, written $[\![w]\!]$, which is defined by the following recursive function:

$$\begin{aligned} [\![\epsilon]\!] &= (0,0) \\ [\![\binom{b_1}{b_2} \cdot w]\!] &= (2 * m + b_1,\ 2 * n + b_2) \\ &\qquad \textit{where } (m,\ n) = [\![w]\!] \end{aligned}$$

For example:

$$[\![\binom{0}{1}\binom{0}{0}\binom{1}{1}\binom{0}{1}]\!] = (4,\ 13)$$

$$[\![\binom{0}{1}\binom{1}{1}\binom{0}{0}\binom{1}{0}\binom{1}{0}]\!] = (26,\ 3)$$

$$[\![\binom{1}{1}\binom{0}{1}\binom{0}{1}\binom{0}{0}\binom{0}{1}]\!] = (1,\ 23)$$

For each of the following, give a finite state automaton that recognises the (encoding of the) language and has at most 3 states. For this problem, your automata should *not* accept the empty word.

i. $\{w \mid [\![w]\!] = (n, m) \wedge n \leq m\}$

ii. $\{w \mid [\![w]\!] = (n, m) \wedge m = 2 * m\}$

[8 marks]

(d) Let Σ be an alphabet and suppose $M_1 = (Q_1, \Sigma, \delta_1, p_1, F_1)$ and $M_2 = (Q_2, \Sigma, \delta_2, p_2, F_2)$ are deterministic finite state automata. Describe a finite state automaton that recognises the following language:

$$\{a_1 b_1 a_2 b_2 \cdots a_n b_n \mid n \geq 0 \wedge a_1 a_2 \cdots a_n \in L(M_1) \wedge b_1 b_2 \cdots b_n \in L(M_2)\}$$

(i.e. those words of even length where concatenating the letters at even numbered positions yields a word accepted by M_1 and concatenating the letters at odd-numbered positions yields a word accepted by M_2). *[6 marks]*

(e) Let $\#_0(w)$ be the number of '0' letters in word w and $\#_1(w)$ be the number of '1' letters in word w. For example, $\#_0(01101) = 2$ and $\#_1(01101) = 3$.

Prove that the language $\{w \in \{0,1\}^* \mid \#_0(w) = \#_1(w)\}$ is not regular.

[6 marks]

(cont.)

(f) Given two languages A and B over a common alphabet Σ, define:

$$A \triangleleft B := \{w \in \Sigma^* \mid \exists v.\, wv \in A \wedge v \in B\}$$

Suppose A is regular. Show that there is a finite automaton recognising $A \triangleleft B$ (irrespective of whether or not B is regular).

[8 marks]

Q2. This question is about the While language.

(a) i. Consider the following While program.

```
n := 1
r := 0
while (n <= x) {
  n := 2 * n
  r := r + 1
}
```

Describe the 7th configuration in its execution trace starting in initial state $[\mathrm{x} \mapsto 4]$.

ii. Write a program in While that always terminates in a state where variable z has value x^y, where x (resp. y) is the initial value of variable x (resp. y). You may assume that we only run this program in initial states where y is non-negative.

iii. Consider the nth Fibonacci number $\mathrm{fib}(n)$, defined inductively as follows.

$$\mathrm{fib}(n) = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n = 1 \\ \mathrm{fib}(n-2) + \mathrm{fib}(n-1) & \text{otherwise} \end{cases}$$

Write a While program that computes the nth Fibonacci number for any given $n \in \mathbb{N}$. Your program should read the value of n from variable n in its initial state (you can assume it is always non-negative), and write the value of $\mathrm{fib}(n)$ into variable result in its final state. You can use as many additional variables as desired.

[15 marks]

(b) This question part is about the Abstract Machine's bytecode language and about translating from While to the Abstract Machine.

Consider the program c_{Q2b} shown in Figure 1. It is written in the Abstract Machine's bytecode language.

For $m, n \in \mathbb{Z}$, define the initial state $s(m,n)$ as $s(m,n) = \begin{bmatrix} \mathrm{a} \mapsto m; \\ \mathrm{b} \mapsto n \end{bmatrix}$.

i. What are the values of variables q and r in the final state of the execution of c_{Q2b} in $s(5,4)$ (if it exists)?

ii. What are the values of variables q and r in the final state of the execution of c_{Q2b} in $s(100,-10)$ (if it exists)?

```
PUSH 0; STORE q;
LOOP(LOAD a; LOAD b; LE;
   , PUSH 1; LOAD q; ADD; STORE q;
     LOAD b; LOAD a; SUB; STORE a;);
FETCH a; STORE r;
```

Figure 1: The Abstract Machine program c_{Q2b}.

iii. Under what conditions on the values of a and b does the adversary terminate when run in state $s(a, b)$?

iv. More generally, given a and b such that $0 \leq a$ and $0 < b$, what is the value of variables q and r in the final state of the execution of c_{Q2b}?

v. Euclid's algorithm computes the greatest common divisor of two positive integers a and b. We define it below as a function $\gcd(a, b)$, where $b \mod a$ is the remainder in the Euclidean division of b by a.

$$\gcd(a,b) = \begin{cases} b & \text{if } a = 0 \\ \gcd(b,a) & \text{if } b < a \\ \gcd(b \mod a, a) & \text{otherwise} \end{cases}$$

Produce a program in the Abstract Machine's bytecode language which implements Euclid's algorithm. You may assume that inputs are given in variables a and b, and that your program will only be run in states that give them positive values. Place the output in variable result.

[15 marks]

(c) This question part is a more advanced compilation problem. You probably should not attempt it until you are convinced you cannot do more on the rest of the exam.

We consider a knock-off Abstract Machine (the Knock-Off Machine, below) that can only hold two elements on its stack. We give (part of) its structural operational semantics as a transition relation $\triangleright_{\text{KO}}$ (Figure 2).

$$\begin{array}{rll}
\langle \texttt{PUSH}\ n; c, e, s\rangle & \triangleright_{\text{KO}} \langle c, n : e, s\rangle & \text{if } |e| \leq 1 \\
\langle \texttt{ADD}; c, z_1 : z_2, s\rangle & \triangleright_{\text{KO}} \langle c, z_1 + z_2, s\rangle & \text{if } z_1, z_2 \in \mathbb{Z} \\
\langle \texttt{MUL}; c, z_1 : z_2, s\rangle & \triangleright_{\text{KO}} \langle c, z_1 * z_2, s\rangle & \text{if } z_1, z_2 \in \mathbb{Z} \\
\langle \texttt{SUB}; c, z_1 : z_2, s\rangle & \triangleright_{\text{KO}} \langle c, z_1 - z_2, s\rangle & \text{if } z_1, z_2 \in \mathbb{Z} \\
\langle \texttt{FETCH}\ \mathsf{x}; c, e, s\rangle & \triangleright_{\text{KO}} \langle c, s(\mathsf{x}) : e, s\rangle & \text{if } |e| \leq 1 \\
\langle \texttt{STORE}\ \mathsf{x}; c, z : e, s\rangle & \triangleright_{\text{KO}} \langle c, e, s[\mathsf{x} \mapsto z]\rangle & \text{if } z \in \mathbb{Z}
\end{array}$$

Figure 2: Semantics for the Knock-Off Machine's arithmetic instructions

Define and argue correct a translation function $\mathcal{CA}_{\text{KO}}[\![\cdot]\!]$ from Imp's *arithmetic expressions only* to the Abstract Machine's bytecode language.

When the source expression a has defined semantics $\mathcal{A}[\![a]\!](s)$ in a given state s, executing the translated program $C\mathcal{A}_{\mathsf{KO}}[\![a]\!]$ with the empty evaluation stack in state s should yield a configuration with an empty program, a stack containing exactly one element whose value is $\mathcal{A}[\![a]\!](s)$, and a state that must be larger than s in the sense that it is defined and equal to s wherever s is defined, and may also give values to variables not defined in s.

Your translation function may rely on an environment and does not need to be optimal. In fact, simpler is better.

[5 marks]

Q3. This question is about computability.

(a) A perfect square is a number which is of the form n^2 for some $n \in \mathbb{N}$. Show that the set

$$U = \{n \in \mathbb{N} \mid \ n \text{ is a perfect square } \}$$

is decidable.

[2 marks]

(b) Let $f : A \to B$ and $g : B \to C$. Show that if $g \circ f : A \to C$ is injective, then so is f.

[2 marks]

(c) Let us say that a state σ is 2022-*bounded* just if, for all variables x, $-2022 \leq \sigma(x) \leq 2022$.

One of the following two predicates P and Q is semi-decidable and the other is not. Determine which one is semi-decidable and justify your answer.

$P = \{\gamma(S) \mid \text{for all 2022-bounded } \sigma,\ S \text{ terminates when started in } \sigma\}$

$Q = \{\gamma(S) \mid \text{there is 2022-bounded } \sigma \text{ and } S \text{ does not terminate when started in } \sigma\}$

[8 marks]

(d) Show that the following predicate is undecidable:

$$P = \{\phi(\gamma(S_1), \gamma(S_2)) \mid \text{for all } n \in \mathbb{N}\text{: } [\![S_1]\!]_x(n) \simeq 1 \text{ iff } [\![S_2]\!]_x(n) \simeq k \text{ where } k \neq 1 \ \}$$

[8 marks]

END OF PAPER