THE UNIVERSITY OF BRITISH COLUMBIA

CPSC 421/501: FINAL EXAMINATION - PART 1 - December 22, 2021

Full Name: xLASTNAMEx CS Ugrad ID: xFIRSTNAMEx

Signature: ______________________ UBC Student #: [illegible]

Important notes about this examination

1. You have 60 minutes to complete this part of the exam.
2. The exam is closed book, and students may bring in two double-sided sheets of notes (8.5" x 11").
3. Good luck!

Student Conduct during Examinations

1. Each examination candidate must be prepared to produce, upon the request of the invigilator or examiner, his or her UBCcard for identification.
2. Examination candidates are not permitted to ask questions of the examiners or invigilators, except in cases of supposed errors or ambiguities in examination questions, illegible or missing material, or the like.
3. No examination candidate shall be permitted to enter the examination room after the expiration of one-half hour from the scheduled starting time, or to leave during the first half hour of the examination. Should the examination run forty-five (45) minutes or less, no examination candidate shall be permitted to enter the examination room once the examination has begun.
4. Examination candidates must conduct themselves honestly and in accordance with established rules for a given examination, which will be articulated by the examiner or invigilator prior to the examination commencing. Should dishonest behaviour be observed by the examiner(s) or invigilator(s), pleas of accident or forgetfulness shall not be received.
5. Examination candidates suspected of any of the following, or any other similar practices, may be immediately dismissed from the examination by the examiner/invigilator, and may be subject to disciplinary action:
   i. speaking or communicating with other examination candidates, unless otherwise authorized;
   ii. purposely exposing written papers to the view of other examination candidates or imaging devices;
   iii. purposely viewing the written papers of other examination candidates;
   iv. using or having visible at the place of writing any books, papers or other memory aid devices other than those authorized by the examiner(s); and,
   v. using or operating electronic devices including but not limited to telephones, calculators, computers, or similar devices other than those authorized by the examiner(s)—(electronic devices other than those authorized by the examiner(s) must be completely powered down if present at the place of writing).
6. Examination candidates must not destroy or damage any examination material, must hand in all examination papers, and must not take any examination material from the examination room without permission of the examiner or invigilator.
7. Notwithstanding the above, for any mode of examination that does not fall into the traditional, paper-based method, examination candidates shall adhere to any special rules for conduct as established and articulated by the examiner.
8. Examination candidates must follow any additional examination rules or directions communicated by the examiner(s) or invigilator(s).

1 236413 048929

0. Identification

Please make sure that the following is your 5-character ugrad email id:

xFIRSTNAMEx

Your answer to each problem should be written on its page; if needed, you can use the back side of the page as well.

1. (10 POINTS, 2 POINTS PER CORRECT T/F ANSWER — NO PENALTY FOR INCORRECT RESPONSES)

Circle either T for true, or F for false, for each of the statements below.

(a) Let L be any language over $\{a, b\}$ that is recognized by some NFA with 30 states. Then there must exist a DFA that recognizes L with at most 900 states. T **(F)**

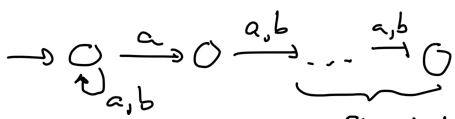
(b) The language $\{a^{(3n+7)} \mid n \in \mathbb{N}\} = \{a^{10}, a^{13}, a^{16}, a^{19}, \ldots\}$ is regular. **(T)** F

(c) The language $\{a^{(n^2)} \mid n \in \mathbb{N}\} = \{a, a^4, a^9, a^{16}, \ldots\}$ is regular. T **(F)**

(d) The set of all regular languages over $\{a, b\}$ is countably infinite. **(T)** F

(e) There exists a surjection from $\mathbb{N}$ to $\{a, b\}^*$. **(T)** F

(a) The language $L = \{ w \in \{a,b\} \mid$ the 29^{th} to last symbol of w is $a\}$ can be recognized by an NFA of 30 states



, but by Myhill-Nerode, as seen in class/homework, requires 2^{29} states in a DFA.

(b) L is of period 3, hence regular

(c) Since $(n+1)^2 - n^2 = 2n+1$, which $\to \infty$ as $n \to \infty$, the gap between successive word lengths $\to \infty$; hence L is not periodic, hence not regular

(d) Since $\{a,b\}^*$ is countably infinite, the set of languages over $\{a,b\}$, i.e. Power($\{a,b\}^*$), is uncountable

(e) $\{a,b\}^*$ is countable (by writing $\{a,b\}^* = \{\varepsilon\} \cup \{a,b\}^1 \cup \{a,b\}^2 \cup \cdots$, each $\{a,b\}^n$ being a finite set).

Note: Some years, such as 2023, the language

$$C_k = \{ w \in \{a,b\} \mid \text{the } k^{th} \text{ last symbol of } w \text{ is "a"} \}$$

is only discussed in class and homework for specific values of k. Hence, in 2023, you should know that C_2 requires 4 states and C_3 requires 8. Hence there are variants of (1a) that you should be able to answer.

Note: The fact that ACCEPT-SOME-INPUT (e.g. class Sept 20, 2023) and $\{ \langle p \rangle \mid p \text{ is a Python program accepting at least 2 inputs} \}$ (e.g. Homework 3, 2023) are both recognizable required us to know that for any alphabet, Σ, the elements of Σ^* can be listed as

$$i_1, i_2, i_3, \ldots$$

Hence in years that we cover this topic, you should know that there is a bijection

$$\mathbb{N} \to \Sigma^* .$$

Hence, in 2023 you would certainly be able to answer (1e). However, we often discuss "countable sets" in CPSC 421/501, which are sets, S, for which there is a surjection

$$\mathbb{N} \to S .$$

Hence in years where we discuss "countable sets", I would expect you to be used to questions like (1e).

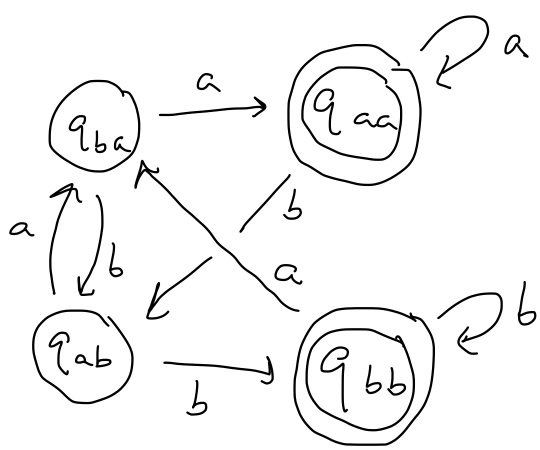
2. (15 POINTS)

Let $L = \{a,b\}^*\{aa, bb\}$, i.e., the language of words over $\Sigma = \{a,b\}$ that end in either aa or bb.

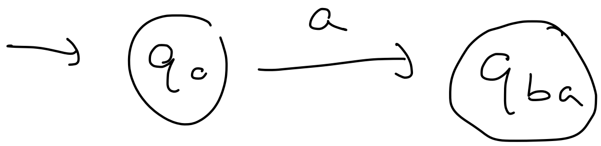
(1) Write a formal description of a DFA (either by a diagram, table of δ values, or listing every value of a DFA) that recognizes L.

(2) Use the Myhill-Nerode theorem (and possibly your solution to the previous part) to determine the minimum number of states of a DFA needed to recoginze L (over the alphabet $\Sigma = \{a,b\}$).

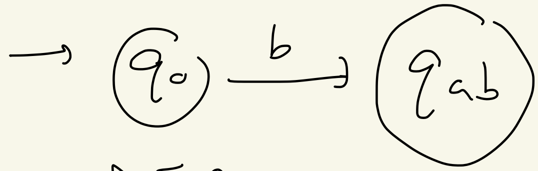
(2.1) We remember the last 2 characters read; $q_{ba}, q_{aa}, q_{ab}, q_{bb}$ correspond, respectively, to the last 2 charecters ba, aa, ab, bb; this gives us



The start of the DFA has an initial state q_0 for ε; since membership is determined by only the last two symbols, and the "b" in a word ending in "ba" doesn't affect whether or not the word ends in "aa" or "bb," we can take

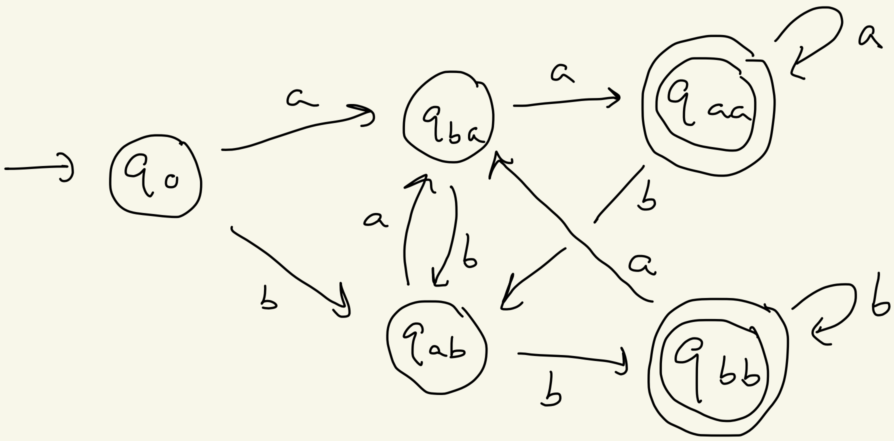


,

and similarly 

.

Hence we get the DFA



NOTE: (2.1) does not ask for an explanation; however, if you make a mistake in your DFA, then an explanation can give you more marks.

(2.2) $AccFut_L(w)$ depends only on the last 2 symbols in w, and if u has length ≥ 2, then whether or not $wu \in L$ does not depend on w. Hence for any $w \in \{a,b\}^*$,

$$AccFut_L(w) = L \cup \{u \mid u \text{ has length } 0,1 \text{ and } wu \in L\}.$$

Hence if $|w| \geq 2$, $\mathrm{AccFut}_L(w)$ is one of

$\mathrm{AccFut}_L(aa)$, $\mathrm{AccFut}_L(ab)$, $\mathrm{AccFut}_L(ba)$, $\mathrm{AccFut}(bb)$.

Also, by the above

$$\mathrm{AccFut}_L(ab) = L \cup \{b\}$$
$$\mathrm{AccFut}_L(aa) = L \cup \{\varepsilon, a\}$$
$$\mathrm{AccFut}_L(ba) = L \cup \{a\}$$
$$\mathrm{AccFut}_L(bb) = L \cup \{\varepsilon, b\}$$

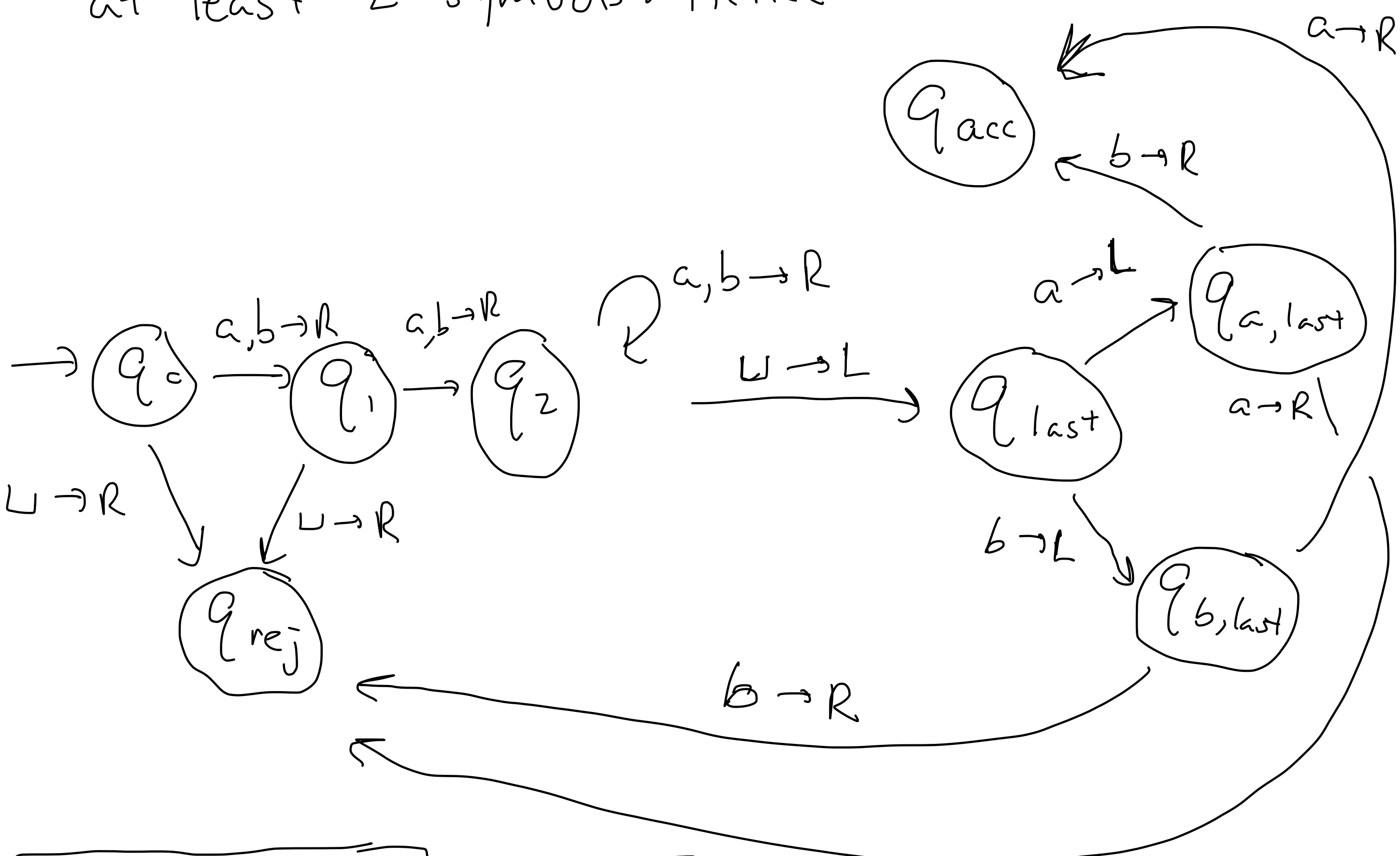
since L consists only of words of length ≥ 2, and $\{b\}, \{\varepsilon, a\}, \{a\}, \{\varepsilon, b\}$ are visibly distinct, these sets are different. Since $\mathrm{AccFut}_L(\varepsilon) = L$, which contains no words of length ≤ 1, $\mathrm{AccFut}_L(\varepsilon)$ is different from the above four AccFut_L.

Hence any DFA accepting L must have at least 5 states. Hence, by Myhill-Nerode and (2.1), the minimum number of states is 5.

3. (10 POINTS)

Let $L = \{a, b\}^*\{ab, ba\}$, i.e., the language of words over $\Sigma = \{a, b\}$ that end in either ab or ba. Write a formal description of a Turing machine that recognizes L, and give a **brief explanation (one to three sentences)** of how your Turing machine works; make sure you clearly indicate $Q, \delta, \Gamma, q_0, q_{\text{accept}}, q_{\text{reject}}$ and **all values** of δ.

We can scan the input, moving R (right) until we reach a blank symbol ⊔; then we move L twice, remembering the last symbol and hence deciding if the word is in L or not. Also: we need to make sure that the word has at least 2 symbols. Hence



Q = the 8 states above

Σ = {a, b}

Γ = {a, b, ⊔}

δ, q_0, q_{acc}, q_{rej} as indicated above

4. (10 POINTS)

Let

$$L = \{a^n b^{(n^2)} \mid n \in \mathbb{N}\} = \{ab, a^2b^4, a^3b^9, \ldots\}.$$

Prove that L is not regular. [You may use the Myhill-Nerode theorem, or any other facts covered in our course.]

The shortest word in $\mathrm{AccFut}_L(a^m)$ for $m \in \mathbb{N}$ is b^{m^2} (other words have some extra a's, followed by more than m^2 b's).

Hence for $m \in \mathbb{N}$, $\mathrm{AccFut}_L(a^m)$ are all different. Hence a is not regular.

Alternative (in years when we talk about the substitution method): mapping $a \mapsto \varepsilon$ and $b \mapsto b$, we see that if L is regular, then so is the image of L under this map, which is $\{b^{(n^2)}\}$. Since this language is not regular (see Problem 1(c)), L is not regular.