THE UNIVERSITY OF BRITISH COLUMBIA
CPSC 421: MIDTERM EXAMINATION – October 30, 2019

Full Name: xLASTNAMEx   Exam ID: xFIRSTNAMEx

Signature: ______________________   UBC Student #: [illegible]

Important notes about this examination

1. You have **50** minutes to complete this examination.
2. One two sided (8.5” x 11”) sheet of notes is allowed.
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. No questions will be answered in this exam. If you see text you feel is ambiguous, make a reasonable assumption, write it down, and proceed to answer the question.
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).

Please do not write in this space:

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)

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

If L is regular, then L^* is regular. T F

Solution: True: this is a standard closure property of regular languages (class notes 09-25 and [Sip], Section 1.2).

The set of languages over the alphabet $\{a, b\}$ is countable. T F

Solution: False: covered in the article "Self-Referencing, Uncountability, and Uncomputability in CPSC 421" (also, class notes of 09-11).

The set of Turing machines, M, such that for some $q, \gamma \in \mathbb{N}$, the state set of M is $\{1, \ldots, q\}$ and the tape alphabet of M is $\{1, \ldots, \gamma\}$, is countable. T F

Solution: True: we have seen that any such Turing machine can be described as a finite string over the alphabet $\{0, \ldots, 9, \#\}$ (class notes of 10-11).

The union of two nonregular languages is nonregular. T F

Solution: False: if L is any nonregular language over an alphabet Σ, then also the complement of L, $\Sigma^* \setminus L$, is nonregular; but their union of L and its complement is all of Σ^*.

If L is decided by some 5-tape Turning machine, then it is decided by some 1-tape Turing machine. T F

Solution: True: See dicussion of multi-tape Turing machines, starting 10-07 or [Sip], Section 3.2.

2. (10 POINTS)

Let $\Sigma = \{a, b\}$, and let L be the language over Σ given by

$$L = \{s \in \Sigma^* | \ s \text{ contains } ba \text{ as a substring}\}.$$

Describe a Turing machine that decides L, and explain how it works; specify Q and δ either by (1) a list of values or (2) a state diagram. What is your work tape Γ? Clearly indicate which state is your accept state, which is your reject state, and which is your initial state.

Solution: [We will use $q_{\text{init}}, q_{\text{acc}}, q_{\text{rej}}$, respectively as the initial, accept, and reject states.] Our algorithms will always move to the tape head to the right; so we have $\Gamma = \{a, b, \mathit{blank}\}$ (so it is irrelevant which symbols we write in the tape cells). The string has ba as a substring iff the string contains a b, and somewhere after the first b in the string there is an a. Hence our algorithm works as follows:

Phase 1: We scan the input, left to right, waiting to see a b; hence we remain in the intial state, q_{init} until we see a b (reject if we reach a blank without seeing a b), so

$$\delta(q_{\text{init}}, a) = (q_{\text{init}}, \mathit{any}, R), \quad \delta(q_{\text{init}}, \mathit{blank}) = q_{\text{rej}}$$

where *any* is any element of Γ, and where we do not write the tape symbol and R, L tape direction once we reach q_{rej} or q_{acc}.

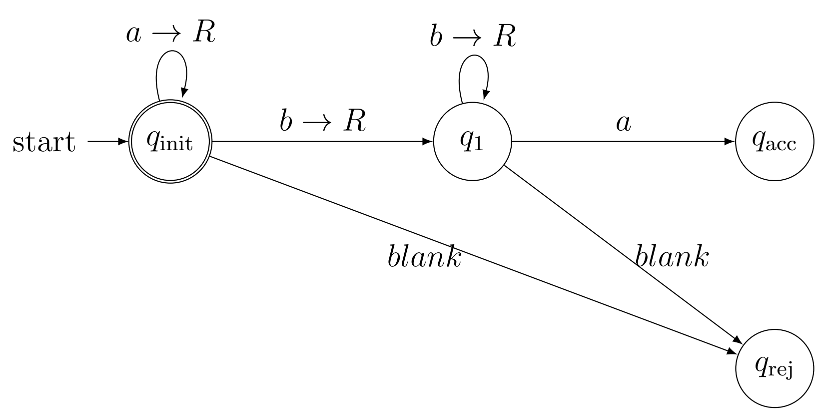
Phase 2: Once we see a b we transition to a state q_1; then we wait in this state until we see an a (whereupon we accept), rejecting if we first reach a *blank* without seeing an a; so

$$\delta(q_{\text{init}}, b) = (q_1, \mathit{any}, R),$$
$$\delta(q_1, b) = (q_1, \mathit{any}, R), \quad \delta(q_1, a) = q_{\text{acc}}, \quad \delta(q_1, \mathit{blank}) = q_{\text{rej}}.$$

We therefore have

$$Q = \{q_{\text{init}}, q_{\text{acc}}, q_{\text{rej}}, q_1\}.$$

As a diagram, this looks like:



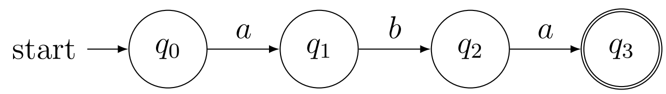
3. (10 POINTS)

Give an NFA recognizing the language $(aba)^*(abaa)^*$; i.e., give a state diagram for the NFA and explain how your NFA works.

Solution: The following NFA recognizes the language $\{aba\}$:

$$\delta(q_0, a) = q_1, \quad \delta(q_1, b) = q_2, \quad \delta(q_2, a) = q_3,$$

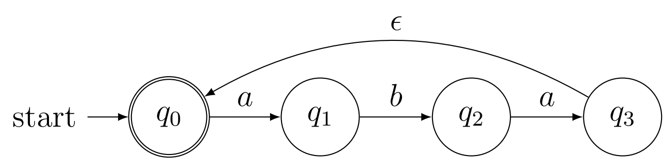
with q_0 initial and q_3 the unique final state. As a diagram this is:



To recognize $(aba)^*$ we add an ϵ transition from all final states (i.e., accepting states) to the initial state and make the initial state accepting (it is probably simplest to make the initial state the only accepting state, although you can leave in any of the other accepting states); hence we add

$$\delta(q_3, \epsilon) = q_0$$

and make q_0 a final state; we will now remove q_3 as a final state, but you could leave it in. This gives the diagram:



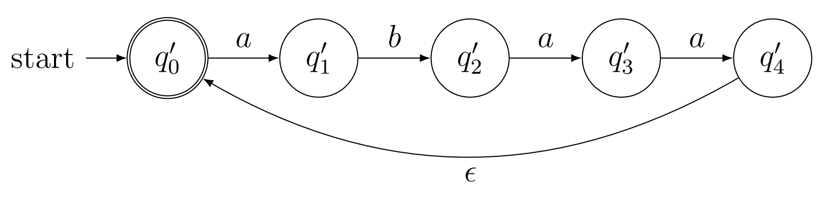
Similarly we recognize $abaa$ via

$$\delta(q_0', a) = q_1', \quad \delta(q_1', b) = q_2', \quad \delta(q_2', a) = q_3', \quad \delta(q_3', a) = q_4',$$

(with q_0' as initial and q_3' as the unique final state), and recognize $(abaa)^*$ adding

$$\delta(q_4', \epsilon) = q_0'$$

and make q_0' a final state (we can remove q_3' as a final state). As a diagram this looks like.

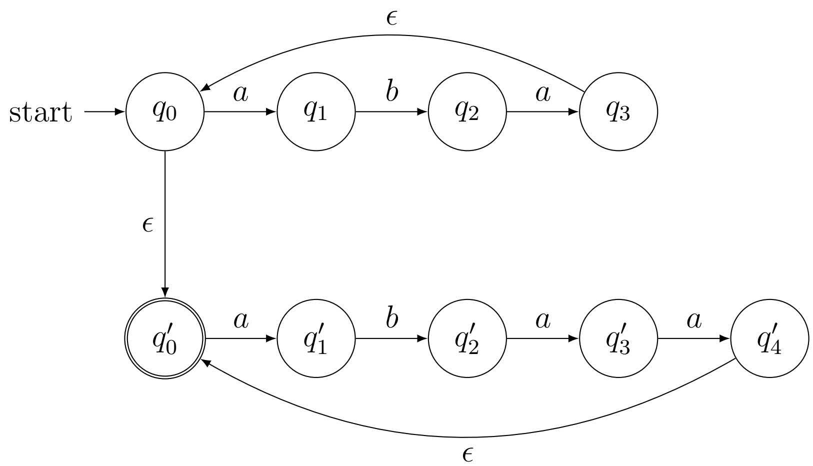


To recognize $(aba)^*(abaa)^*$ we first run the NFA for $(aba)^*$ and then (non-deterministically) make an ϵ transition to the NFA for $(abaa)^*$ from all of the accepting states for the $(aba)^*$ NFA. Since q_0 is the only accepting state for the $(aba)^*$ NFA, we add

$$\delta(q_0, \epsilon) = q_0'.$$

Hence the initial state is q_0, the only accepting state is q_0' (but you could leave any of q_3, q_4', q_0 as accepting).

This yields the NFA with the diagram:


ϵ
start
q_0
a
q_1
b
q_2
a
q_3
ϵ
q'_0
a
q'_1
b
q'_2
a
q'_3
a
q'_4
ϵ

4. (10 POINTS)

Let $\Sigma = \{a, b, c\}$. Use the Myhill-Nerode theorem to show that the language

$$L = \{s \in \Sigma^* \mid \text{Exactly half of the symbols in } s \text{ are } c\text{'s}\}$$

is nonregular, by showing that there are an infinite number of values of $\mathrm{AccFut}_L(s)$ where s varies over Σ^*.

Solution: For any $n \in \mathbb{N}$, the shortest string in $\mathrm{AccFut}_L(a^n)$ is c^n, since a^n has no c's and we need as many c's as a's and b's to be in L. It follows that as n ranges over all of $\mathbb{N}$, all the sets $\mathrm{AccFut}_L(a^n)$ are different. Hence there are an infinite number of accepting futures for L, and hence, by the Myhill-Nerode theorem, L is not regular.