

THE UNIVERSITY OF BRITISH COLUMBIA
CPSC 421/501: MIDTERM EXAMINATION – November 14, 2025

Last Name: _____

First Name: _____

Signature: _____

UBC Student #: _____

Important notes about this examination

1. You have 45 minutes to write this examination.
2. You may use a pencil to write your solutions, although a very light pencil might be harder to read after scanning.
3. No textbooks or electronic devices are permitted. We permit a “cheat-sheet” consisting of 1 two-sided 8.5x11 sheet of notes, handwritten or typewritten.
4. Answer all the questions in the exam.
5. 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).



0. IDENTIFICATION

Please make sure that the following is your 8-character Student ID:

Student ID:

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

: Some brief notes:

Regular expressions consist of $\emptyset$, any element of Σ , or ϵ plus the operations $\cup$, $\circ$, and $*$ (and parenthesis when needed/wanted).

Recognizable means Python-Recognizable (this is equivalent to Turing-recognizable, but the midterm won't cover Turing machines).

A language recognizable by a DFA is a "regular", and will not be called "recognizable" (even though "DFA-recognizable" would be correct).

1. QUESTION 1. (12 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. In these questions, $\Sigma = \Sigma_{\text{ASCII}}$ denotes the ASCII alphabet, $L_1, L_2 \subset \Sigma^*$, $L_1 \setminus L_2$ denotes the set difference of L_1 and L_2 .

If L_1, L_2 are finite languages, then $|L_1 \circ L_2| = |L_2 \circ L_1|$.

T ☒ F

E.g. $L_1 = \{ba, baa\}$, $L_2 = \{a, aa\}$

If L_1 and L_2 are both recognizable, then $L_1 \setminus L_2$ is recognizable.

T ☒ F

E.g. $L_1 = \Sigma^*$, $L_2 = \text{HALTING}_{\text{python}}$

$\text{Power}(\Sigma^*)$ is countable.

T ☒ F

Σ^* is countably infinite, so $\text{Power}(\Sigma^*)$ is uncountable

The set of Python-recognizable languages over Σ is countable.

☒ T F

Each such language is recognized by a Python program in Σ^* , and Σ^* is countable

The set of functions $\Sigma^* \rightarrow \{1, 2\}$ is countably infinite.

T ☒ F

This is in bijection with $\text{Power}(\Sigma^*)$

The set of functions $\{1, 2\} \rightarrow \Sigma^*$ is countably infinite.

☒ T F

This is a bijection with $\Sigma^* \times \Sigma^*$

and hence a bijection with $\mathbb{N} \times \mathbb{N}$,
which we have seen is countably infinite

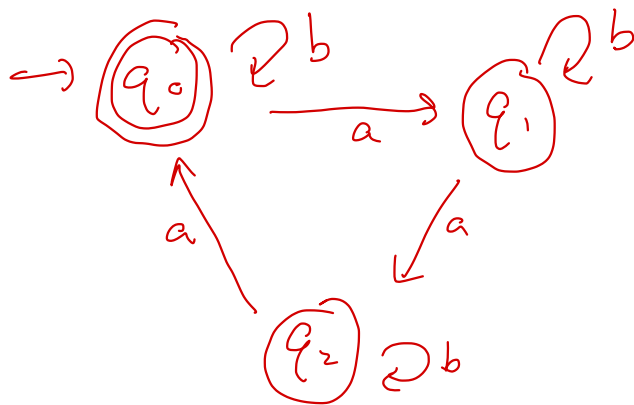
2. QUESTION 2 (4 POINTS)

Let $\Sigma = \{a, b\}$, and let $L \subset \Sigma^*$ be the set of strings whose total number of a 's is an integer divisible by 3 (including 0). Hence

$$L = \{\epsilon, aaa, aaab, abba, abaa, baaa, aaabb, \dots\}.$$

Build a DFA that recognizes L and briefly explain how your DFA works.

$$L = \{ \epsilon, b, bb, aaa, bbb, aaab, \dots \}$$



Strings land in q_0 if they lie in L ; for $i=0,1,2$
 " " " q_i if they contain m a 's with
 $m \bmod 3 = i$.

Hence on a "b" we remain the same state

On an "a" we move +1 modulo 3, so

$$q_0 \rightarrow q_1, \quad q_1 \rightarrow q_2, \quad q_2 \rightarrow q_0$$

We accept a string if it lands in q_0 , since then (and only then) is the # of a 's mod 3 equal to 0 (i.e. # of a 's is divisible by 3).

3. QUESTION 3 (8 POINTS)

Let $\Sigma = \{a\}$ and $\Sigma' = \{a, b\}$. Let

$$L = \{a^m \mid m \geq 2\} = \{a^2, a^3, a^4, \dots\}.$$

- (1) Using the Myhill-Nerode Theorem, find the minimum number of states in a DFA over Σ needed to recognize L .

$$\text{Accfut}_L(\varepsilon) = L \quad \leftarrow \text{min length string: } a^2$$

$$\text{Accfut}_L(a) = \{a, a^2, a^3, \dots\} \quad \leftarrow \dots \dots \dots : a$$

$$\text{Accfut}_L(a^2) = \{\varepsilon, a, a^2, \dots\} = \{a\}^* \quad \leftarrow \dots \dots \dots : \varepsilon$$

and for all $m \geq 3$

$$\text{Accfut}_L(a^m) = \{a\}^* = \text{Accfut}_L(a^2)$$

Since min length string in $\text{Accfut}_L(a^i)$ for $i = 0, 1, 2$ is a^{2-i} , we have 3 distinct accepting futures.

Also $\text{Accfut}_L(a^m)$ with $m \geq 2$ all the same. So min # states in a DFA is 3

- (2) Using the Myhill-Nerode Theorem, find the minimum number of states in a DFA over Σ' needed to recognize L .

$\text{Accfut}_L(a^i)$, $i = 0, 1, \dots$ are same, giving 3 accepting futures.

If s has 1 or more b 's, then

$\text{Accfut}(s) = \emptyset$. Hence total # different accepting

futures over $\{a, b\}$ is 4. So min number

states over $\{a, b\}$ is 4.

4. QUESTION 4 (5 POINTS)

Let VALID-PYTHON be the set of ASCII strings that are valid Python programs. L be the language

$$L = \{p \in \text{VALID-PYTHON} \mid p \text{ accepts at least one input}\},$$

i.e.,

$$L = \{p \in \text{VALID-PYTHON} \mid \exists i \in \text{ASCII}^* \text{ such that } p \text{ accepts } i\}.$$

Show that L is recognizable.

List ASCII^* as $i_1, i_2, i_3, \dots$

Consider the algorithm:

Phase 0: check that p is a valid program, (reject if not)

Phase 1: run p for 1 step on i_1 (with a universal program)

Phase 2: " " " 2 steps on each of i_1, i_2
 $\vdots$

Phase k : " " " k " " " " " $i_1, \dots, i_k$
 $\vdots$

If any input is accepted on some phase, accept p .

Then: if p accepts some string, say string i_m after n steps, then we accept in phase k where $k = \max(n, m)$ (which is finite).

Each phase runs in a finite amount of time.

Otherwise our algorithm accepts p iff $p \in L$.