

Name (first last)

SID

- This exam is **closed book, closed notes** and is designed to take 45 minutes
- Turn off you cell phones
- Write legibly. What can't be read will not be graded
- Good luck!

1	<input type="text"/>	/ 16
2	<input type="text"/>	/ 16
3	<input type="text"/>	/ 18
4	<input type="text"/>	/ 18
5	<input type="text"/>	/ 16
6	<input type="text"/>	/ 16
Total	<input type="text"/>	/100

1. (16 points: 2 points if correct, 1 point if unanswered, 0 points if wrong.)

Mark by true or false each of the following (no need to prove)

A language cannot be both regular and context free at the same time True False

The family of regular languages is closed under union True False

The grammar $G = (\{S\}, \{a\}, S, \{S \rightarrow aS \mid Sa \mid a\})$ is regular True False

The grammar $G = (\{S\}, \{a, b\}, S, \{S \rightarrow aaS \mid bbS \mid b\})$ is not regular True False

Given a string w and two regular languages L_1 and L_2 there exists an algorithm to determine whether $w \in (L_1 \cap L_2)$ True False

The string $(\emptyset(\lambda\emptyset)^*)^*$ is a regular expression True False

The language associated with the regular expression True False

$$(1 + 01)^*(0 + \lambda)$$

is the set of the binary strings containing no occurrences of 00

The grammar $G = (\{S, A, B\}, \{a\}, S, \{S \rightarrow AB, A \rightarrow a \mid BA, B \rightarrow \lambda \mid AB\})$ is not in Chomsky normal form True False

2. (16 points)

Write the formal definition of a grammar in Chomsky normal form (CNF).

Answer: A context-free grammar $G = (V, T, S, P)$ is said to be *Chomsky normal form* if all the productions in P have the form $A \rightarrow BC$ or $C \rightarrow \mathbf{a}$ where $A, B, C \in V$ and $\mathbf{a} \in T$.

Write the formal definition of the language associated with a regular expression.

Answer: The language $L(r)$ associated with any regular expression r is defined by the following rules.

- $L(\emptyset) = \emptyset$
- $L(\lambda) = \{\lambda\}$
- $L(a) = \{a\}$ where $a \in \Sigma$

Let r_1 and r_2 be two regular expressions, then

- $L(r_1 + r_2) = L(r_1) \cup L(r_2)$
- $L(r_1 \cdot r_2) = L(r_1)L(r_2)$
- $L(r_1^*) = L(r_1)^*$
- $L((r_1)) = L(r_1)$

3. (18 points)

Let L_1, L_2, \dots, L_9 be the following languages over $\Sigma = \{a, b\}$

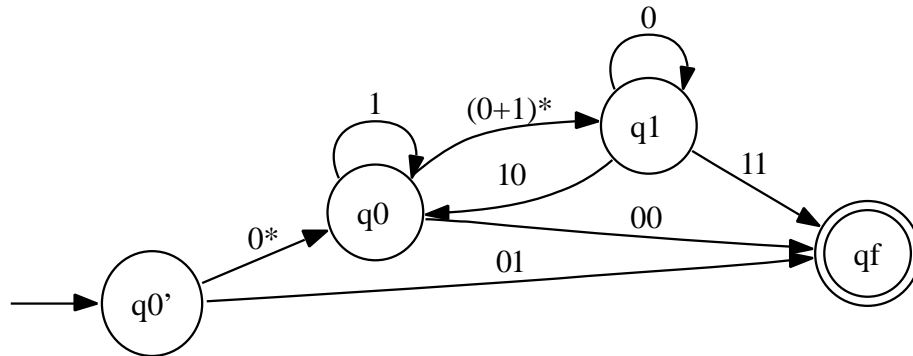
$$\begin{aligned}
 L_1 &= \{a^i b^j : j \geq i \geq 0\} \\
 L_2 &= \{a^i b^j : i \geq j \geq 0\} \\
 L_3 &= \{a^i b^i a^j b^j : i, j \geq 0\} \\
 L_4 &= \{a^i b^j a^j b^i : i, j \geq 0\} \\
 L_5 &= \{ww^R : w \in \Sigma^*\} \\
 L_6 &= \{a^i b^j a^i : i, j \geq 0\} \\
 L_7 &= \{a^{i+j} b^i a^j : i, j \geq 0\} \\
 L_8 &= \{a^i b^j a^{i+j} : i, j \geq 0\} \\
 L_9 &= \{a^i b^{i+j} a^j : i, j \geq 0\}
 \end{aligned}$$

In the table below you are given five context-free grammars. For each grammar choose **one** language among L_1, \dots, L_9 generated by this grammar (or write “none” if it does not generate any of these languages).

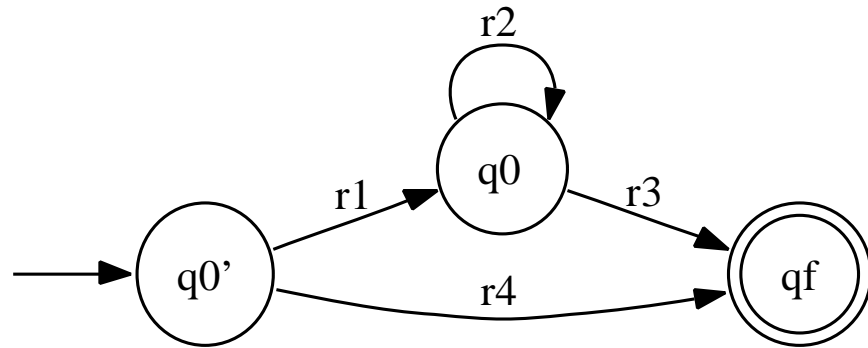
<i>Grammar</i>	<i>Language</i>
$S \rightarrow YX$ $X \rightarrow bXa \mid \lambda$ $Y \rightarrow aYb \mid \lambda$	L_9
$S \rightarrow aSb \mid X$ $X \rightarrow bX \mid \lambda$	L_1
$S \rightarrow aSa \mid X$ $X \rightarrow bXa \mid \lambda$	L_8
$S \rightarrow aSb \mid X$ $X \rightarrow bXa \mid \lambda$	L_4
$S \rightarrow bSb \mid aSa \mid \lambda$	L_5

4. (18 points)

The figure below shows an automaton obtained in the process of converting an NFA to a regular expression. Complete the process by eliminating first q_1 and then q_0 .

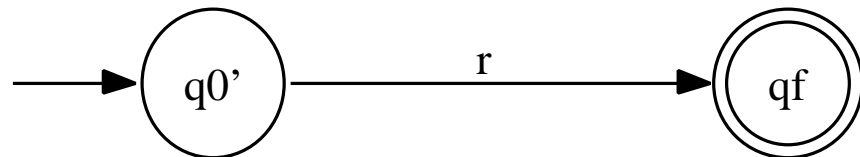


Eliminate q_1 :



Where $r_1 = 0^*$, $r_2 = 1 + (0 + 1)^*0^*10$, $r_3 = (0 + 1)^*0^*11 + 00$, and $r_4 = 01$.

Eliminate q_0 :



Where $r = r_1r_2^*r_3 + r_4$.

5. (16 points)

Prove that the language

$$L = \{a^n b^k : n \geq 0, k \geq n\}$$

is not regular. We begin the proof assuming that the opponent has chosen a constant m .

- What string do you choose for $w \in L$, such that $|w| \geq m$?
Answer: $w = a^m b^m$ (other choices may be also correct).
- Suppose that the adversary decomposes the string $w = xyz$ such that $|xy| \leq m$ and $y \neq \lambda$. What value of i do you choose to create a string $w_i = xy^i z$ such that w_i is not in L ?
Answer: $i = 2$ (other choices may be also correct).
- Given your choice of i , explain **briefly** why $xy^i z$ does not belong to L .
Answer: The string y is only composed by a 's, so that we can write $y = a^l, l > 0$. By choosing $i = 2$, we get $w_2 = a^{m+l} b^m$ which does not belong to L because the number of a 's is now greater than the number of b 's.

6. (16 points)

Remove λ -productions, unit-productions, and useless productions, from the following grammar. Show each step of the simplification work.

$$\begin{aligned} S &\rightarrow AC | Ca \\ A &\rightarrow BS \\ B &\rightarrow SA \\ C &\rightarrow aS | \lambda \end{aligned}$$

Answer: Remove λ -productions. Since $V_N = \{C\}$, we have

$$\begin{aligned} S &\rightarrow AC | A | Ca | a \\ A &\rightarrow BS \\ B &\rightarrow SA \\ C &\rightarrow aS \end{aligned}$$

Remove unit-productions. Since $S \Rightarrow^* A$, we have

$$\begin{aligned} S &\rightarrow AC | BS | Ca | a \\ A &\rightarrow BS \\ B &\rightarrow SA \\ C &\rightarrow aS \end{aligned}$$

Remove useless productions. First, reachability. It is easy to verify that you can reach all the variables from S . Next, productivity. We have that $V_1 = \{S, C\}$ (in other words, A and B are non-productive) hence

$$\begin{aligned} S &\rightarrow Ca | a \\ C &\rightarrow aS \end{aligned}$$